

Result	Query	Score	Match	Length	ID	Description
1	3363	100.0	3363	24	ARS56553	CDNA encod
2	3366.8	99.8	3495	22	AK52348	Human poly
3	3348.4	99.6	4303	22	AAD05336	Human secr
4	3348.4	99.6	4303	24	ABL9787	Human poly
5	3347.2	99.5	4397	24	ABN8354	Human KIAA
6	3346.8	99.5	4286	24	ARS55559	Full-length
7	3344.8	99.5	3348	22	AAH76201	Human drug
8	3323.4	98.8	3906	21	AAZ64969	Membrane-b





1381	TGTCAGCGTGCTGAGTATCACAGACGGCGTGTGACGAGCTGGGACAGAAAGTGTGGCATGTGTGTG 1444	Db
1441	GAGGACGCCACCGGGAGCTGAAGAGCTGCATAAAGTGCAAAGGGCCCCCATGCGGCTGGGCGGCG 1500	QY
1441	GAGGACGCCACCGGGAGCTGAAGAGCTGCATAAAGTGCAAAGGGCCCCCATGCGGCTGGGCGGCG 1500	Db
1501	AGCAGAGCCCTCTCCAAACCTCGTGGCCAGTAGTACTACGGGCAAGGCGACGAGGCGCTGCACC 1560	QY
1501	AGCAGAGCCCTCTCCAAACCTCGTGGCCAGTAGTACTACGGGCAAGGCGACGAGGCGCTGCACC 1560	Db
1561	TGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGGGACGCCGGAAAAAACTCTTCAAGAAAG 1620	QY
1561	TGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGGGACGCCGGAAAAAACTCTTCAAGAAAG 1620	Db
1621	AAGTACAAGGCGAGCTATGTCGCCAGTGCCTCCATCGGCTCAGTGGCCATCGAGGTGGAC 1680	QY
1621	AAGTACAAGGCGAGCTATGTCGCCAGTGCCTCCATCGGCTCAGTGGCCATCGAGGTGGAC 1680	Db
1681	GGCAGGGTGATACCAAGCTAGGCGCTGGGTGATGCGCGCCAGCGCCGAAACCTCACCAAGCGG 1740	QY
1681	GGCAGGGTGATACCAAGCTAGGCGCTGGGTGATGCGCGCCAGCGCCGAAACCTCACCAAGCGG 1740	Db
1741	CATCGGCGAGGGGCCCTGAGGACCAAGATGACAAGATGGTGGGACTTCAAGTGGCACT 1800	QY
1741	CATCGGCGAGGGGCCCTGAGGACCAAGATGACAAGATGGTGGGACTTCAAGTGGCACT 1800	Db
1801	GGAGGCGTCTCCGACTACTCAGCGCGCAACCCCATTAAGTGGACACATCGGTGCTACATC 1860	QY
1801	GGAGGCGTCTCCGACTACTCAGCGCGCAACCCCATTAAGTGGACACATCGGTGCTACATC 1860	Db
1861	CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGCGCTGGAAA 1920	QY
1861	CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGCGCTGGAAA 1920	Db
1921	GACCAAGGCTGCACATCGACACGAGATGAAACCTGCGAGNACAAATTAAGAAGCTTG 1980	QY
1921	GACCAAGGCTGCACATCGACACGAGATGAAACCTGCGAGNACAAATTAAGAAGCTTG 1980	Db
1981	AGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCGGCCAGAAAGATGTCACTGTCACAAAATC 2040	QY
1981	AGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCGGCCAGAAAGATGTCACTGTCACAAAATC 2040	Db
2041	AGTACCAACCCAGACAAAAGGCGCGCTCAAGCAAGAGGCTCAGTGTGCATCCTTTTC 2100	QY
2041	AGTACCAACCCAGACAAAAGGCGCGCTCAAGCAAGAGGCTCAGTGTGCATCCTTTTC 2100	Db
2101	AGGAAGGGCTGCAAGAGAAAGGACAAGGTGTGGCTGTGCGGAGCAGAAAGCGCAAGAG 2160	QY
2101	AGGAAGGGCTGCAAGAGAAAGGACAAGGTGTGGCTGTGCGGAGCAGAAAGCGCAAGAG 2160	Db
2161	AAACTCGCAAGCTGCTCAAGCGCTTCGAGAACCAACGACACGTGACGATGCCAGGCTTC 2220	QY
2161	AAACTCGCAAGCTGCTCAAGCGCTTCGAGAACCAACGACACGTGACGATGCCAGGCTTC 2220	Db
2221	ACGTGCTTCAACCAACGACACGAGCTGGCAGACGGCGCTTTCTGGACATCTGGGGCTTC 2280	QY
2221	ACGTGCTTCAACCAACGACACGAGCTGGCAGACGGCGCTTTCTGGACATCTGGGGCTTC 2280	Db
2281	TTCTGTGCTTCGACCAAGCGCCAAACAATAACACGCTACTGTGTGATGAGGACCATCAATGAG 2340	QY
2281	TTCTGTGCTTCGACCAAGCGCCAAACAATAACACGCTACTGTGTGATGAGGACCATCAATGAG 2340	Db
2341	ACTCACAAAATTCCTCTCTGTGAATTTGMACTGGGCTTCCTAGAGTACTTTGATCTCAAC 2400	QY
2341	ACTCACAAAATTCCTCTCTGTGAATTTGMACTGGGCTTCCTAGAGTACTTTGATCTCAAC 2400	Db
2401	ACAGACCCCTTACCAGCTGATGAATGACAGTGAAACACACTGGAAGGGATGTCTCTCAACCCAG 2460	QY
2401	ACAGACCCCTTACCAGCTGATGAATGACAGTGAAACACACTGGAAGGGATGTCTCTCAACCCAG 2460	Db
2461	CTACAGCTACAGCTCAATGGAGCTGAGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520	QY
2461	CTACAGCTACAGCTCAATGGAGCTGAGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520	Db

Qy	2521	ACTCGAACAATGAGCACTGGGACCTTAAAGATGGAGGAAGCTATGAGCGAAATACAGGCAGTTT	2580
Db	2521	ACTCGAACAATGAGCACTGGGACCTTAAAGATGGAGGAAGCTATGAGCGAAATACAGGCAGTTT	2580
Qy	2581	CAGCGTCGAAGTGGCCAGAAAATGAAGAGACCTTCTTCCAAATCAGCTGGGACAACTGTGG	2640
Db	2581	CAGCGTCGAAGTGGCCAGAAAATGAAGAGACCTTCTTCCAAATCAGCTGGGACAACTGTGG	2640
Qy	2641	GAAGGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCCAAAACAATAGAGGCAATCACC	2700
Db	2641	GAAGGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCCAAAACAATAGAGGCAATCACC	2700
Qy	2701	TGACTGCACAGGCAATGAAAAACAATGTGGTGCATTTCCAGACAGACCTGTGCTATTGGCC	2760
Db	2701	TGACTGCACAGGCAATGAAAAACAATGTGGTGCATTTCCAGACAGACCTGTGCTATTGGCC	2760
Qy	2761	AGGAGGCTCGAGAAAGCAAGCACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC	2820
Db	2761	AGGAGGCTCGAGAAAGCAAGCACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC	2820
Qy	2821	CAGCAGGACAGAGATAAATTTCAGGAAGTCCATTTTGGCCCTGCTTTTCTGTTTGGATT	2880
Db	2821	CAGCAGGACAGAGATAAATTTCAGGAAGTCCATTTTGGCCCTGCTTTTCTGTTTGGATT	2880
Qy	2881	TACCTCACCAGCTGCACAAAATGCATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC	2940
Db	2881	TACCTCACCAGCTGCACAAAATGCATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC	2940
Qy	2941	CCAGAGCTCACAAGGAAAACCGAGAGACGAGCGAGAGAGATTTTCCTTGGAAATTTCT	3000
Db	2941	CCAGAGCTCACAAGGAAAACCGAGAGACGAGCGAGAGAGATTTTCCTTGGAAATTTCT	3000
Qy	3001	CCAGAGGCGAAAAGTCAATCGGAATTTTAAATCATAGGGGAAAAGCAGTCCCTGTCTTAAA	3060
Db	3001	CCAGAGGCGAAAAGTCAATCGGAATTTTAAATCATAGGGGAAAAGCAGTCCCTGTCTTAAA	3060
Qy	3061	TCCTCTATTCTTTTGGTTTGTCCAAAAGGAACTTAAGAAGCAGGACAGAGCAACGCT	3120
Db	3061	TCCTCTATTCTTTTGGTTTGTCCAAAAGGAACTTAAGAAGCAGGACAGAGCAACGCT	3120
Qy	3121	GGAGAGGCTGAAAACAGTCAGAGAGCGTTTGCAATGAGTCAGTACGACAAAAGAGATGA	3180
Db	3121	GGAGAGGCTGAAAACAGTCAGAGAGCGTTTGCAATGAGTCAGTACGACAAAAGAGATGA	3180
Qy	3181	CATTACCTAGCACTATAAACCCCTGGTGGCTCTGAAGAACTGCCCTTCAATGTATATAT	3240
Db	3181	CATTACCTAGCACTATAAACCCCTGGTGGCTCTGAAGAACTGCCCTTCAATGTATATAT	3240
Qy	3241	GTGACTATTACATGTAATCAACATGGAACTTTTAGGGGAACCTTAATAGAAATCCCAA	3300
Db	3241	GTGACTATTACATGTAATCAACATGGAACTTTTAGGGGAACCTTAATAGAAATCCCAA	3300
Qy	3301	TTTTTCAGGAGTGGTGTGTCAATAAAGCGCTCTGTGGCGAGTGTAAAAAGAAAAA	3360
Db	3301	TTTTTCAGGAGTGGTGTGTCAATAAAGCGCTCTGTGGCGAGTGTAAAAAGAAAAA	3360
Qy	3361	AAA 3363	
Db	3361	AAA 3363	

## RESULT 2

AAK523348  
ID AAK52348 standard; cDNA; 3495 BP.

AAK52348;

DT 06-NOV-2001 (first entry)

XX  
DE Human polynucleotide SEQ ID NO 893.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy; KW

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZH;

XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX P-PSDB; AAM79215.

XX Nucleic acids encoding polypeptides with cytokine-like activities,  
XX useful in diagnosis and gene therapy -

XX Claim 1; Page 2939-2943; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation.

XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
XX (AAM80020) are omitted as the relevant pages from the sequence listing  
XX were missing at the time of publication.

XX Sequence 3495 BP; 958 A; 976 C; 924 G; 637 T; 0 other;  
XX Query Match 99.8%; Score 3356.8; DB 22; Length 3495;  
XX Best Local Similarity 99.9%; Pred. No. 0;  
XX Matches 3358; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTCCGCGATCCCAAAGAGACACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 60

DB 136 GACTCCGCGATCCCAAAGAGACACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 195

QY 61 CTCGTGTGTGCTGTGTCGCGAAGTGTCTTCTCCCTGCTGGTGGAGCTGGCTTC 120

DB 196 CTCGTGTGTGCTGTGTCGCGAAGTGTCTTCTCCCTGCTGGTGGAGCTGGCTTC 255

QY 121 CTGTGCGACACACCGCTGAAAGGAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 180

DB 256 CTGTGCGACACACCGCTGAAAGGAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 315

QY 181 ATCATCTGTGTGTCGAGGACGACAGATGTGGAGCTGGGTTCATGAGGTGATGAC 240

DB 316 ATCATCTGTGTGTCGAGGACGACAGATGTGGAGCTGGGTTCATGAGGTGATGAC 375

QY 241 AAGACCGGCGCATCATGTGAGCAGGGCGGCGGCGACATTCATCAACGCTTCGTGACCA 300

DB 376 AAGACCGGCGCATCATGTGAGCAGGGCGGCGGCGACATTCATCAACGCTTCGTGACCA 435

QY 301 CCCATGTGTGCTGCTCAGCTCCTCCTCATCTGCAAGTACGTCCCAACACCAAC 360

DB 436 CCCATGTGTGCTGCTCAGCTCCTCCTCATCTCACCAGTACGTCCCAACACCAAC 495

QY 361 ACCTACACCAACATGAGAACTGCTCCTCGCCCTCCTGGCAGGACACAGCAGAGAGCGC 420

DB 496 ACCTACACCAACATGAGAACTGCTCCTCGCCCTCCTGGCAGGACACAGCAGAGAGCGC 555

QY 421 ACCTTTCGCTGTGCTCAATAGCACTGGCTACCGGACAGCTTTCCTTGGGAAGTATCTT 480

DB 556 ACCTTTCGCTGTGCTCAATAGCACTGGCTACCGGACAGCTTTCCTTGGGAAGTATCTT 615

QY 481 AATGAATAACAACGGCTCCTACGTGCCACCGGCTGGAAGAGTGGTGGGACTCCTTAAA 540

DB 616 AATGAATAACAACGGCTCCTACGTGCCACCGGCTGGAAGAGTGGTGGGACTCCTTAAA 675

QY 541 AACTCCCGCTTTTATAACTACAGCTGTGTGCGAAACGGGGTGAAGAAAGCAGCGCTCC 600

DB 676 AACTCCCGCTTTTATAACTACAGCTGTGTGCGAAACGGGGTGAAGAAAGCAGCGCTCC 735

QY 601 GACTACTCCAAAGGATTAACCTCAGAGACTCATACCAATGACAGCGTGAGCTTCTTCGCG 660

DB 736 GACTACTCCAAAGGATTAACCTCAGAGACTCATACCAATGACAGCGTGAGCTTCTTCGCG 795

QY 661 ACCTCCAAAGAGATGTACCCGACAGCGCTCCTCATGTGCTCATGCCATGACGCGCTCC 720

DB 796 ACCTCCAAAGAGATGTACCCGACAGCGCTCCTCATGTGCTCATGCCATGACGCGCTCC 855

QY 721 CACGCGCTGAGGATTCAGCCCAACATATTCAGCGCTCTTCCCAACGCGATCTCAGCAC 780

DB 856 CACGCGCTGAGGATTCAGCCCAACATATTCAGCGCTCTTCCCAACGCGATCTCAGCAC 915

QY 781 ATCAGCGCGAGCTCAACTACGCGCCCAACACCGGACAAACACTGATCATGCGCTACACG 840

DB 916 ATCAGCGCGAGCTCAACTACGCGCCCAACACCGGACAAACACTGATCATGCGCTACACG 975

QY 841 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGAGCGCTTCAG 900

DB 976 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGAGCGCTTCAG 1035

QY 901 ACCCTCATGTGCTGGAGAGTCCATCGAGAGAGCTTCAACATGCTGGTGGAGAGCGGC 960

DB 1036 ACCCTCATGTGCTGGAGAGTCCATCGAGAGAGCTTCAACATGCTGGTGGAGAGCGGC 1095

QY 961 GAGCTGGAACAACAGTACATGATATACCGCGACACCGGTTACCATCGGCGAGTTT 1020

DB 1096 GAGCTGGAACAACAGTACATGATATACCGCGACACCGGTTACCATCGGCGAGTTT 1155

QY 1021 GGCCTGTGTAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1080

DB 1156 GGCCTGTGTAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1215

QY 1081 AGGGCCCCCAACGCTGGAAGCGGCTGTCTGAATCCCACTGCTCCTCAACATGAGCTG 1140

DB 1216 AGGGCCCCCAACGCTGGAAGCGGCTGTCTGAATCCCACTGCTCCTCAACATGAGCTG 1275

QY 1141 GCGCCCACTCTGGAACATGAGGCTGACATACCTCGGATATGGAGCGGAATCC 1200

DB 1276 GCGCCCACTCTGGAACATGAGGCTGACATACCTCGGATATGGAGCGGAATCC 1335

QY 1201 ATCTCTCAAGCTGCTGGACACAGGAGCGGCTGATCGGTTTCACTTTGAAAAGAGATG 1260

DB 1336 ATCTCTCAAGCTGCTGGACACAGGAGCGGCTGATCGGTTTCACTTTGAAAAGAGATG 1395

QY 1261 AGGGCTGCGCGGACTCCTTCTGCTGAGAGAGGAGGAGCTGTCTACACAGAGAGACAT 1320

DB 1396 AGGGCTGCGCGGACTCCTTCTGCTGAGAGAGGAGGAGGAGCTGTCTACACAGAGAGACAT 1455

QY 1321 GACAGAGTGGAGCGCCAGGAGGAGAACTTTCTGCGCAAGTACACAGCGTGTGAGGAGCTG 1380

[illegible]

### RESULT 3

AAD05336

AAD05336  
 ID AAD05336 standard; cDNA; 4303 BP.

XX  
DT

AC AAD

AC XX FAD

DT 17-

17

XX DE HUM

KW  
 XX  
 H1E

foetal abnormality; developmental abnormality; haematopoietic disorder;  
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
inflammation; allergy; neurological disorder; Alzheimer's disease;  
Parkinson's disease; cognitive disorder; schizophrenia; aschma;  
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
cardiovascular disorder; angioinetic disorder; kidney disorder;  
gastrointestinal disorder; pregnancy-related disorder;  
endocrine disorder; infection; wound healing; vulnery;  
cell culture; chemotaxis; food additive; gene therapy;  
binding partner identification; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 222..2834

XX /product= "Human secreted protein"

XX /transl\_except= [pos:1821..1823, aa:Xaa]

XX /note= "Xaa equals any of the twenty naturally occurring

XX L-amino acids"

XX sig\_peptide 222..293

XX mat\_peptide 294..2831

XX /tag= a

XX /tag= b

XX /tag= c

XX /product= "Mature human secreted protein"

XX WO200134626-A1.

XX 17-MAY-2001.

XX 01-NOV-2000; 2000WO-US30045.

XX 05-NOV-1999; 99US-0163581.

XX 30-JUN-2000; 2000US-0215133.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;

XX WPI: 2001-308778/32.

XX P-PSDB; AAE01471.

XX New nucleic acid molecules encoding 28 human secreted proteins for  
diagnosing, preventing, treating or ameliorating medical conditions and  
used as food additives or preservatives -  
Claim 1; Page 450-452; 562pp; English.

XX AAD05300-AA05379 represent cDNAs corresponding to 28 human secreted  
protein genes, and AAE01436-AAE01513 represent the proteins they encode.  
XX AAE01514-AAE01544 represent human secreted protein fragments or variants.  
XX The genes and their secreted proteins are useful for preventing,  
XX treating or ameliorating medical conditions, e.g., by protein or gene  
XX therapy. Pathological conditions can be diagnosed by determining the  
XX amount of the new protein in a sample or by determining the presence of  
XX mutations in the new genes. Specific uses are described for each of the  
XX 28 genes, based on the tissues in which they are most highly expressed,  
XX and include developing products for the diagnosis or treatment of  
XX proliferative disorders, cancer, tumours, foetal and developmental  
XX abnormalities, haematopoietic disorders, diseases of the immune system,  
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
XX allergies, neurological disorders (e.g., Alzheimer's disease,  
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
XX cardiovascular disorders, angioinetic disorders, kidney disorders,  
XX gastrointestinal disorders, pregnancy-related disorders, endocrine  
XX disorders, and infections. The proteins can also be used to aid wound  
XX healing and epithelial cell proliferation, to prevent skin aging due to  
XX sunburn, to maintain organs before transplantation, for supporting cell  
XX culture of primary tissues, to regenerate tissues, to identify their  
XX cognate ligands or binding partners, and in chemotaxis, and can be used  
XX as a food additive or preservative to modify storage properties.  
XX Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and  
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
CC immunosorbent assay (ELISA). The present sequence represents a human  
CC secreted protein-encoding cDNA of the invention.

XX Sequence 4303 BP; 1195 A; 1142 C; 1073 G; 889 T; 4 other;

XX Query Match 99.6%; Score 3348.4; DB 22; Length 4303;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 3349; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTCCCGCATCCCAAAAGAGACACAGATCAGCAAAAAGAGAGATGGGCCCCCGGAGC 60  
DB 177 GACTCCCGCATCCCAAAAGAGACACAGATCAGCAAAAAGAGAGATGGGCCCCCGGAGC 236  
QY 61 CTCGTGCTGTCTTCTGCTCCGCAACTGTCTTCTCCCTGCTGGGTGGAGTCTGGGCTTC 120  
DB 237 CTCGTGCTGTCTTCTGCTCCGCAACTGTCTTCTCCCTGCTGGGTGGAGTCTGGGCTTC 296  
QY 121 CTGTGCGACCAACCGCTGAAAGGCGAGTTTCAGAGGACCGCAGGAAACATCCGCCCAAC 180  
DB 297 CTGTGCGACCAACCGCTGAAAGGCGAGTTTCAGAGGACCGCAGGAAACATCCGCCCAAC 356  
QY 181 ATCATCTCTGCTGTGACGAGACGACGAGATGTGGAGCTGGGTTCATCAGTGTATGAAC 240  
DB 357 ATCATCTCTGCTGTGACGAGACGACGAGATGTGGAGCTGGGTTCATCAGTGTATGAAC 416  
QY 241 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTTCATCAAGCCTTCGTGACCAACA 300  
DB 417 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTTCATCAAGCCTTCGTGACCAACA 476  
QY 301 CCCATGTCTGCCCTCACGCTCCCTCATCTCTCACTGCGCAAGTACGTCCACAACCAAC 360  
DB 477 CCCATGTCTGCCCTCACGCTCCCTCATCTCTCACTGCGCAAGTACGTCCACAACCAAC 536  
QY 361 ACCTACCAACAATGAGAACTGCTCTCGCCCTCTCTGCGGACGACAGCAGAGAGCGGC 420  
DB 537 ACCTACCAACAATGAGAACTGCTCTCGCCCTCTCTGCGGACGACAGCAGAGAGCGGC 596  
QY 421 ACCTTGGCGGTACTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 480  
DB 597 ACCTTGGCGGTACTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 656  
QY 481 AATGAATACAAACGGCTCTCTACGTCACCCCGCTCGGAGAGTGGGTGGAGTCTCTTAAA 540  
DB 657 AATGAATACAAACGGCTCTCTACGTCACCCCGCTCGGAGAGTGGGTGGAGTCTCTTAAA 716  
QY 541 AACTCCCGCTTTTAACTAACAAGCTGTGTGGAAACGGGTGAAAGAAAGACAGCGCTCC 600  
DB 717 AACTCCCGCTTTTAACTAACAAGCTGTGTGGAAACGGGTGAAAGAAAGACAGCGCTCC 776  
QY 601 GACTACTCCAAAGGATTACTCTACAGACCTCATCAACAATGACAGCTGAGCTTCTTCGCG 660  
DB 777 GACTACTCCAAAGGATTACTCTACAGACCTCATCAACAATGACAGCTGAGCTTCTTCGCG 836  
QY 661 AGCTCAAGAAAGATGACCCGACAGGCGAGTCTCTATGGTTCATCAGCATCAGCGCCCTC 720  
DB 837 AGCTCAAGAAAGATGACCCGACAGGCGAGTCTCTATGGTTCATCAGCATCAGCGCCCTC 896  
QY 721 CAGCGCCCTGAGGATTACGCCCCCAATATTTCAGCCCTCTTCCCAACAGCATCTCAGCAC 780  
DB 897 CAGCGCCCTGAGGATTACGCCCCCAATATTTCAGCCCTCTTCCCAACAGCATCTCAGCAC 956  
QY 781 ATCAGCCCGAGTCAACAATCAGCCGCAACCCCGGCAAAACACTGGATCATCGCTTACAG 840  
DB 957 ATCAGCCCGAGTCAACAATCAGCCGCAACCCCGGCAAAACACTGGATCATCGCTTACAG 1016  
QY 841 GGGCCCATGAAGCCCATCCCATGGAATTCACCAATGCTCCAGGGAAGCGTTGAG 900  
DB 1017 GGGCCCATGAAGCCCATCCCATGGAATTCACCAATGCTCCAGGGAAGCGTTGAG 1076  
QY 901 ACCTCATCTCGGTGGAGCGACTTCCATGGAGAGGATTTACAAATGCTGTTGAGACGGGC 960

1077 AGCTTCACTGCTGGAGCGACTTCATGAGAGAGATTATCAACATGCTGGTTGAGACGGC 1136  
961 GAGCTGGCAACACAGTATACGCTATACACCGCGGACCAAGCTTACCAATCGGCGAGTTT 1020  
1137 GAGCTGGCAACACAGTATACGCTATACACCGCGGACCAAGCTTACCAATCGGCGAGTTT 1196  
1021 GGCCTGGTGAAGGAAATCCATGCGCATATGAGTTTGAATCAGGCTCCGTTCTACGTTG 1080  
1197 GGCCTGGTGAAGGAAATCCATGCGCATATGAGTTTGAATCAGGCTCCGTTCTACGTTG 1256  
1081 AGGGGCCCCAAGCTGGAGCGGGCTGTCTGAATCCCACTGCTCCTCAACATTTGACCTG 1140  
1257 AGGGGCCCCAAGCTGGAGCGGGCTGTCTGAATCCCACTGCTCCTCAACATTTGACCTG 1316  
1141 GCGCCCAACATCTCTGGACATTTGAGGCTGGACATACCTGCGGATATGAGCGGAATTC 1200  
1317 GCGCCCAACATCTCTGGACATTTGAGGCTGGACATACCTGCGGATATGAGCGGAATTC 1376  
1201 ATCTCTAAGCTCTGGACACGAGCGGGCTGTAATCGGTTTCACTTGAAGAAAGATG 1260  
1377 ATCTCTAAGCTCTGGACACGAGCGGGCTGTAATCGGTTTCACTTGAAGAAAGATG 1436  
1261 AGGGTCTGGCGGACTCCTTTCTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAAT 1320  
1437 AGGGTCTGGCGGACTCCTTTCTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAAT 1496  
1321 GACAAGTGGACCCCGGAGAGAGAACTTTCTGCCCAAGTACCCAGCGTGTGAAGGACTG 1380  
1497 GACAAGTGGACCCCGGAGAGAACTTTCTGCCCAAGTACCCAGCGTGTGAAGGACTG 1556  
1381 TGTCAGCGTCTGAGTACACAGACGCGTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG 1440  
1557 TGTCAGCGTCTGAGTACACAGACGCGTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG 1616  
1441 GAGGACGCCACCGGGGAAGCTGAAGCTGCATAGTGCATAGGCGGCCCATGCGCTGGGCGGC 1500  
1617 GAGGACGCCACCGGGGAAGCTGAAGCTGCATAGTGCATAGGCGGCCCATGCGCTGGGCGGC 1676  
1501 AGCAGAGCCCTCTCAACCTCTGTCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 1560  
1677 AGCAGAGCCCTCTCAACCTCTGTCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 1736  
1561 TGTCAGCGGGGACTACAGCTCAGCTGCGCGGACCGGACCGGAAAAAACTCTTCAAGAG 1620  
1737 TGTCAGCGGGGACTACAGCTCAGCTGCGCGGACCGGACCGGAAAAAACTCTTCAAGAG 1796  
1621 AAGTACAGGCGAGTATGTCCGAGTGCCTCCATCCGCTCAGTGGCCCATCGAGTGGAC 1680  
1797 AAGTACAGGCGAGTATGTCCGAGTGCCTCCATCCGCTCAGTGGCCCATCGAGTGGAC 1856  
1681 GGCAGGGTGTACAGCTAGGCGCTGGGTGATGCGGCCCGGACCGGAAAACTCAACAGCGG 1740  
1857 GGCAGGGTGTACAGCTAGGCGCTGGGTGATGCGGCCCGGACCGGAAAACTCAACAGCGG 1916  
1741 CACTGGCAGGGGCGGCTGAGGACCAAGATGATGATGAGGAGTGTGGGACTTCAGTGGCACT 1800  
1917 CACTGGCAGGGGCGGCTGAGGACCAAGATGATGATGAGGAGTGTGGGACTTCAGTGGCACT 1976  
1801 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCATTAAGTGAACATCGGTGTACATC 1860  
1977 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCATTAAGTGAACATCGGTGTACATC 2036  
1861 CTAGAGAACGACACAGTCCAGTGTGACTTGGACCTGTGTACAAAGTCCCTCAGGCGCTGGAAA 1920  
2037 CTAGAGAACGACACAGTCCAGTGTGACTTGGACCTGTGTACAAAGTCCCTCAGGCGCTGGAAA 2096  
1921 GACCAAGCTGCATCGACACGAGATGGAACCTCTGCGAACAATAATTAAGAACCTG 1980  
2097 GACCAAGCTGCATCGACACGAGATGGAACCTCTGCGAACAATAATTAAGAACCTG 2156  
1981 AGGGAGTCCGAGGTCACCTGAAAGAAAAAGCGCGGAGAGATGTGACTGTACAAAAATC 2040  
2157 AGGGAGTCCGAGGTCACCTGAAAGAAAAAGCGCGGAGAGATGTGACTGTACAAAAATC 2216

2041 AGCTACCAACACCGAGCAAAAGCGCGCTCAAGACACAGAGGCTCCAGTGTGCATCCTTTTC 2100  
2217 AGCTACCAACACCGAGCAAAAGCGCGCTCAAGACACAGAGGCTCCAGTGTGCATCCTTTTC 2276  
2101 AGGAAGGGCTGCAAGAGAAAGCAAGGTGTGGCTGTTCGGGAGCAGAAAGCGCAAGAAG 2160  
2277 AGGAAGGGCTGCAAGAGAAAGCAAGGTGTGGCTGTTCGGGAGCAGAAAGCGCAAGAAG 2236  
2161 AAACCTCCGAAGCTGCTCAAGCGCTGCAAGAAACACACACGCTGCAGCATGCCAGGCTTC 2220  
2337 AAACCTCCGAAGCTGCTCAAGCGCTGCAAGAAACACACACGCTGCAGCATGCCAGGCTTC 2396  
2221 ACCTGCTTCAACCAAGCAACGAGCATCTGGGAGACGCGGCTTTCTGGACATCTGGGCTTC 2280  
2397 ACCTGCTTCAACCAAGCAACGAGCATCTGGGAGACGCGGCTTTCTGGACATCTGGGCTTC 2456  
2281 TTCTGTGCTGCAACGAGCGCCCAACATTAACAGTACTGTGTGATAGGACCATCAATGAG 2340  
2457 TTCTGTGCTGCAACGAGCGCCCAACATTAACAGTACTGTGTGATAGGACCATCAATGAG 2516  
2341 ACTCAAAATTTCTTCTGTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400  
2517 ACTCAAAATTTCTTCTGTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2576  
2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGGACAGGATGCTCTCAACGAG 2460  
2577 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGGACAGGATGCTCTCAACGAG 2636  
2461 CTACAGCTCAGCTCATGAGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGAACCCCGG 2520  
2637 CTACAGCTCAGCTCATGAGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGAACCCCGG 2696  
2521 ACTCGAAAACTGGAACCTGGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 2580  
2697 ACTCGAAAACTGGAACCTGGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 2756  
2581 CAGCGTCGAAAGTGGCGGACAAATGAAGAGACCTTTTCCAAATCACTGGGACAACTGTGG 2640  
2757 CAGCGTCGAAAGTGGCGGACAAATGAAGAGACCTTTTCCAAATCACTGGGACAACTGTGG 2816  
2641 GAAGGCTGGGAAGTTAAGAAACAACAGAGGTGGACCTTCCAAAAACATAGAGGCTACCC 2700  
2817 GAAGGCTGGGAAGTTAAGAAACAACAGAGGTGGACCTTCCAAAAACATAGAGGCTACCC 2876  
2701 TGACTCAGACGCAATGAAGAAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760  
2877 TGACTCAGACGCAATGAAGAAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2936  
2761 AGGAGCCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCTGGAGGATAAC 2820  
2937 AGGAGCCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCTGGAGGATAAC 2996  
2821 CAGCAGGACAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
2997 CAGCAGGACAGATTAACCTTCAGGAGTCCATTTTGGGAAGTCCATTTTGGCTTTGGATTA 3056  
2881 TACCTCAGCAGCTGCAAAAAATGCAATTTTTCGTATCAAAAAAGTCAACCACTAACCCCTCC 2940  
3057 TACCTCAGCAGCTGCAAAAAATGCAATTTTTCGTATCAAAAAAGTCAACCACTAACCCCTCC 3116  
2941 CAGAGAGCTCACAAGAGAAAAAGGAGAGCGGAGAGAGATTTCCTTGGAAAAATTTCT 3000  
3117 CAGAGAGCTCACAAGAGAAAAAGGAGAGCGGAGAGAGATTTCCTTGGAAAAATTTCT 3176  
3001 CCCAAGGGGAAAGTCAATTTGGAAATTTTAAATCATAGGGGAAAAAGCAGTCTGTCTTAAA 3060  
3177 CCCAAGGGGAAAGTCAATTTGGAAATTTTAAATCATAGGGGAAAAAGCAGTCTGTCTTAAA 3236  
3061 TCCTCTTATTTCTTTGGTTTGTCAAGAAAGAACTAAGAAAGAGGACAGAGGCAACT 3120  
3237 TCCTCTTATTTCTTTGGTTTGTCAAGAAAGAACTAAGAAAGAGGACAGAGGCAACT 3296



QY 3121 GGAGAGGCTGAAACAGTGCAGAGACGCTTTGACATGAGTCACTAGCACAAGAGATGA 3180  
Dd |||||  
Dd 3297 GGAGAGGCTGAAACAGTGCAGAGACGCTTTGACATGAGTCACTAGCACAAGAGATGA 3356  
QY 3181 CATTACTAGCATATAAACCTGGTTGGCTCTGAAGAACTGCTTCAATGTATATAT 3240  
Dd |||||  
Dd 3357 CATTACTAGCATATAAACCTGGTTGGCTCTGAAGAACTGCTTCAATGTATATAT 3416  
QY 3241 GTGACTATTACATGTAATCAACATGGAACCTTTTAGGGGAACCTTAATAGAAATCCCA 3300  
Dd |||||  
Dd 3417 GTGACTATTACATGTAATCAACATGGAACCTTTTAGGGGAACCTTAATAGAAATCCCA 3476  
QY 3301 TTTTCAGAGTGGTGGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAA 3352  
Dd |||||  
Dd 3477 TTTTCAGAGTGGTGGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAA 3528

## RESULT 4

ABL90787  
ID ABL90787 standard; cDNA; 4303 BP.  
XX AC ABL90787;  
XX DT 24-MAY-2002 (first entry)  
XX DE Human polynucleotide SEQ ID NO 1349.  
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200190304-A2.  
XX PD 29-NOV-2001.  
XX PF 18-MAY-2001; 2001WO-US16450.  
XX PR 19-MAY-2000; 2000US-205515P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Birse CE, Rosen CA;  
XX DR WPI; 2002-122018/16.  
XX PT P-PSDB; ABB90378.  
XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive, proliferative  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX PS Claim 4; SEQ ID NO 1349; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL9449-ABL90853) and proteins  
CC (ABE89040-ABE90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 4303 BP; 1195 A; 1142 C; 1073 G; 889 T; 4 other;  
Query Match 99.6%; Score 3348.4; DB 24; Length 4303;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3349; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GACTCCCGCATCCCAAGAGCAGCAGATCAGCAAAAGAGAGATGGGCCCCCGCAGC 60  
Dd |||||  
Dd 177 GACTCCCGCATCCCAAGAGCAGCAGATCAGCAAAAGAGAGATGGGCCCCCGCAGC 236  
QY 61 CTGCTGTGTGCTTGTCTGTCGCAACTGTGTCTTCTCCCTGCTGGGTGGAAGCTCGGCCCTTC 120  
Dd |||||  
Dd 237 CTGCTGTGTGCTTGTCTGTCGCAACTGTGTCTTCTCCCTGCTGGGTGGAAGCTCGGCCCTTC 296  
QY 121 CTGTCGACACCCGCTGAAAGGCGAGGTTTCAGAGGAGCCGAGAGATCCGCCCCCAAC 180  
Dd |||||  
Dd 297 CTGTCGACACCCGCTGAAAGGCGAGGTTTCAGAGGAGCCGAGAGATCCGCCCCCAAC 356  
QY 181 ATCATCTGTGTGCTGACGAGCAGCAGGATGTGAGCTGGGTTCATGCGAGGTGATGAAC 240  
Dd |||||  
Dd 357 ATCATCTGTGTGCTGACGAGCAGCAGGATGTGAGCTGGGTTCATGCGAGGTGATGAAC 416  
QY 241 AAGACCCGCGCATCATGAGAGCAGGCGGGGCGCATTTCTATCAACGCCCTTCGTGACCACA 300  
Dd |||||  
Dd 417 AAGACCCGCGCATCATGAGAGCAGGCGGGGCGCATTTCTATCAACGCCCTTCGTGACCACA 476  
QY 301 CCCATGTGTGTCGCCCTCAGCGCTCTCCATCTCTACTGGCAAGTACGTCCCAACCAAC 360  
Dd |||||  
Dd 477 CCCATGTGTGTCGCCCTCAGCGCTCTCCATCTCTACTGGCAAGTACGTCCCAACCAAC 536  
QY 361 ACTTACACCAACATGAGAACTGCTCTCGCCCTCTCTGGCAGGCACAGCAGAGCGCGC 420  
Dd |||||  
Dd 537 ACTTACACCAACATGAGAACTGCTCTCGCCCTCTCTGGCAGGCACAGCAGAGCGCGC 596  
QY 421 ACCTTTCGCGGTGATCTCAATAGCAGCTGGCTACCGGACAGCTTCTTCGGGAGATATCTT 480  
Dd |||||  
Dd 597 ACCTTTCGCGGTGATCTCAATAGCAGCTGGCTACCGGACAGCTTCTTCGGGAGATATCTT 656  
QY 481 AATGAATACAAACGGCTCTCTAGTGCACACCGGCTGGAAGGAGTGGTTCGAGATCTTTAAA 540  
Dd |||||  
Dd 657 AATGAATACAAACGGCTCTCTAGTGCACACCGGCTGGAAGGAGTGGTTCGAGATCTTTAAA 716  
QY 541 AACTCCCGCTTTTATAACTACACGCTGTGTGGAACCGGGTGAAGAGACAGCGCTCC 600  
Dd |||||  
Dd 717 AACTCCCGCTTTTATAACTACACGCTGTGTGGAACCGGGTGAAGAGACAGCGCTCC 776  
QY 601 GACTACTCCAGGATTAACCTCAGAGACTCATCAACCAATGACAGCGTGAAGTCTTCCGC 660  
Dd |||||  
Dd 777 GACTACTCCAGGATTAACCTCAGAGACTCATCAACCAATGACAGCGTGAAGTCTTCCGC 836  
QY 661 AGTCCCAAGAGATGTACCGGCACAGCGCTGTCTCATGTGTATCATGCGCATGAGCGCCC 720  
Dd |||||  
Dd 837 AGTCCCAAGAGATGTACCGGCACAGCGCTGTCTCATGTGTATCATGCGCATGAGCGCCC 896  
QY 721 CACGCGCTGAGGATTCAGCGCCCAACATATTACGCTCTTCCCAACGATCTTCAGCAC 780  
Dd |||||  
Dd 897 CACGCGCTGAGGATTCAGCGCCCAACATATTACGCTCTTCCCAACGATCTTCAGCAC 956  
QY 781 ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAAAACCTGGATCATGCGGTACAGC 840  
Dd |||||  
Dd 957 ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAAAACCTGGATCATGCGGTACAGC 1016  
QY 841 GGGGCCATGAGGCCCATCCATCGGAATTCACCAACATGTCCAGCGGAAGCGCTTCAG 900  
Dd |||||  
Dd 1017 GGGGCCATGAGGCCCATCCATCGGAATTCACCAACATGTCCAGCGGAAGCGCTTCAG 1076  
QY 901 ACCCTCATGTGCGTGGAGCACTCCATCGAGACGATTTTACAACTGCTGGTGGAGACGGCC 960  
Dd |||||  
Dd 1077 ACCCTCATGTGCGTGGAGCACTCCATCGAGACGATTTTACAACTGCTGGTGGAGACGGCC 1136



Qy 961 GAGCTGGACAAACGTAACGTAATACACGGCGGACCAAGGTTTACCAATCGCGCAGTTT 1020  
Db 1137 GAGCTGGACAAACGTAACGTAATACACGGCGGACCAAGGTTTACCAATCGCGCAGTTT 1196  
Qy 1021 GGCTGGTGAAGGGAATCCATGCAATATGAGTTTGAATCAGGGTCCGGTTTACGGT 1080  
Db 1197 GGCTGGTGAAGGGAATCCATGCAATATGAGTTTGAATCAGGGTCCGGTTTACGGT 1256  
Qy 1081 AGGGGCCCAACGTTGAAGCGGCTGTCTGAATCCCGACATCGTCTCAACATTTACCTG 1140  
Db 1257 AGGGGCCCAACGTTGAAGCGGCTGTCTGAATCCCGACATCGTCTCAACATTTACCTG 1316  
Qy 1141 GCCCCHACCATCTCGACATTCAGAGCTGAGAGCTGAGATATGAGCGGGAATCC 1200  
Db 1317 GCCCCHACCATCTCGACATTCAGAGCTGAGAGCTGAGATATGAGCGGGAATCC 1376  
Qy 1201 ATCTCAAGCTGTGGACACGGAGCGGCGGTGAATCGGTTTCACTTTGAAAGAGAGTG 1260  
Db 1377 ATCTCAAGCTGTGGACACGGAGCGGCGGTGAATCGGTTTCACTTTGAAAGAGAGTG 1436  
Qy 1261 AGGCTGTGGCGGAGTCTCTTCTGGTGGAGAGGCGAAGCTGCTACACAGAGAGCAAT 1320  
Db 1437 AGGCTGTGGCGGAGTCTCTTCTGGTGGAGAGGCGAAGCTGCTACACAGAGAGCAAT 1496  
Qy 1321 GACAGGTGACCGCCCAAGGAGGAGACTTTCTGCCAAGTACCAGCGTGTGAGGACCTG 1380  
Db 1497 GACAGGTGACCGCCCAAGGAGGAGACTTTCTGCCAAGTACCAGCGTGTGAGGACCTG 1556  
Qy 1381 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTG 1440  
Db 1557 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTG 1616  
Qy 1441 GAGGACGCCACGGGGAAGCTGAGCTGCATAGTGCAGAGGCCCCATGCGGCTGGCGCGC 1500  
Db 1617 GAGGACGCCACGGGGAAGCTGAGCTGCATAGTGCAGAGGCCCCATGCGGCTGGCGCGC 1676  
Qy 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGCGGCGAGGCGAGCGGCTGCACC 1560  
Db 1677 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGCGGCGAGGCGAGCGGCTGCACC 1736  
Qy 1561 TGTGACAGCGGGGACTACAAAGCTCAGCTTGGCGGAGCGCGGAAAAAACTTTCAAGAA 1620  
Db 1737 TGTGACAGCGGGGACTACAAAGCTCAGCTTGGCGGAGCGCGGAAAAAACTTTCAAGAA 1796  
Qy 1621 AAGTACAGGCGGAGCTATGTCGCGAGTCTCCATCCGCTCAGTGGCCATTCGAGGTGGAC 1680  
Db 1797 AAGTACAGGCGGAGCTATGTCGCGAGTCTCCATCCGCTCAGTGGCCATTCGAGGTGGAC 1856  
Qy 1681 GGCAAGGTGTACCAAGCTAGGCTGGGTGATGCGCGCCAGCGCCCGAAACCTCACCAAGCG 1740  
Db 1857 GGCAAGGTGTACCAAGCTAGGCTGGGTGATGCGCGCCAGCGCCCGAAACCTCACCAAGCG 1916  
Qy 1741 CACTGGCCAGGGGCCCCGAGGACCAAGATGACAGAGTGGTGGGACTTCAGTGGCACT 1800  
Db 1917 CACTGGCCAGGGGCCCCGAGGACCAAGATGACAGAGTGGTGGGACTTCAGTGGCACT 1976  
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACATCGGTCTACATC 1860  
Db 1977 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACATCGGTCTACATC 2036  
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACTGGACTGTGACTGACATCGGTGCTGAGAA 1920  
Db 2037 CTAGAGAACGACACAGTCCAGTGTGACTGGACTGTGACTGACATCGGTGCTGAGAA 2096  
Qy 1921 GACCACAAGCTGCACATCGACACGAGATTGAACCTCGAGAACCAAAATTAAGAACCTG 1980  
Db 2097 GACCACAAGCTGCACATCGACACGAGATTGAACCTCGAGAACCAAAATTAAGAACCTG 2156  
Qy 1981 AGGGAAGTCCGAGGTCACTTGAAGAAAGCGGCGAGAGAAATGTGACTGTGCAAAAATC 2040  
Db 2157 AGGGAAGTCCGAGGTCACTTGAAGAAAGCGGCGAGAGAAATGTGACTGTGCAAAAATC 2216  
Qy 2041 AGCTTACCACCCAGCACAAAGGCGGCTCAAGCACAGAGGCTCCAGTCTGCAATCTTTTC 2100

Db 2217 AGCTTACCACCCAGCACAAAGCGCGCTCAAGCACAGAGGCTCCAGTCTGCAATCTTTC 2276  
Qy 2101 AGGAAGGCGCTGCAAGAGAGAGGCAAGGTGTGGTGTGGGAGAGAGCGCAAG 2160  
Db 2277 AGGAAGGCGCTGCAAGAGAGAGGCAAGGTGTGGTGTGGGAGAGAGCGCAAG 2336  
Qy 2161 AAATCTCGCAAGCTGTCTCAAGCGCTTCAGAGCAACACACAGCTGCAGCATGCCAGGCTC 2220  
Db 2337 AAATCTCGCAAGCTGTCTCAAGCGCTTCAGAGCAACACACAGCTGCAGCATGCCAGGCTC 2396  
Qy 2221 ACCTGCTTACCCACGACCAACCAGCACTGGCAGACGGCGCTTCTGGACATCTGGGCGCT 2280  
Db 2397 ACCTGCTTACCCACGACCAACCAGCACTGGCAGACGGCGCTTCTGGACATCTGGGCGCT 2456  
Qy 2281 TTCTGTGCTGTCACACAGCGCCCAACAATAACACAGTACTGTGTCATGAGAGCAATCAATGAG 2340  
Db 2457 TTCTGTGCTGTCACACAGCGCCCAACAATAACACAGTACTGTGTCATGAGAGCAATCAATGAG 2516  
Qy 2341 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400  
Db 2517 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2576  
Qy 2401 ACAGACCCCTACACAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTCAACAG 2460  
Db 2577 ACAGACCCCTACACAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTCAACAG 2636  
Qy 2461 CTACAGTACAGCTCAGTGGAGCTGAGAGCTGCAAGGTTTCAAGAGTGTAAACCCCGG 2520  
Db 2637 CTACAGTACAGCTCAGTGGAGCTGAGAGCTGCAAGGTTTCAAGAGTGTAAACCCCGG 2696  
Qy 2521 ACTCGAAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGACGTTT 2580  
Db 2697 ACTCGAAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGACGTTT 2756  
Qy 2581 CAGCGTCGAAAGTGGCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Db 2757 CAGCGTCGAAAGTGGCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 2816  
Qy 2641 GAAGCTGGGAGGTTAAGAAACACAGAGTGGACCTCCAAAACATAGAGGATCACC 2700  
Db 2817 GAAGCTGGGAGGTTAAGAAACACAGAGTGGACCTCCAAAACATAGAGGATCACC 2876  
Qy 2701 TGACTGCACAGCGCAATGAAAAACCATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760  
Db 2877 TGACTGCACAGCGCAATGAAAAACCATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2936  
Qy 2761 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAAATGACAGATTCGAGAGGATAAC 2820  
Db 2937 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAAATGACAGATTCGAGAGGATAAC 2996  
Qy 2821 CAGCAGGACGAGATTAACCTTCAGAAAGTCCATTTTGGCCCTCTTTTGGTGGATT 2880  
Db 2997 CAGCAGGACGAGATTAACCTTCAGAAAGTCCATTTTGGCCCTCTTTTGGTGGATT 3056  
Qy 2881 TACCTCACAGCTGCACAAATGCAATTTTCTGATCAAAAAGTCAACATTAACCTCCC 2940  
Db 3057 TACCTCACAGCTGCACAAATGCAATTTTCTGATCAAAAAGTCAACATTAACCTCCC 3116  
Qy 2941 CCAGAGCTCAAAAGGAAACCGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000  
Db 3117 CCAGAGCTCAAAAGGAAACCGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3176  
Qy 3001 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAAGCAGTCTCTTCTAAA 3060  
Db 3177 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAAGCAGTCTCTTCTAAA 3236  
Qy 3061 TCCTCTTATTTCTTTTGTGTCAAAAGAGGAATTAAGAGCAGACAGAGGCAACGT 3120  
Db 3237 TCCTCTTATTTCTTTTGTGTGTCAAAAGAGGAATTAAGAGCAGACAGAGGCAACGT 3296  
Qy 3121 GGAGAGGCTGAAAAAGTGCAGAGACGTTTGCATTAATGAGTCAAGTCAAGAGAGATGA 3180



QY 1021 GGCTGTTGAAGGGAATCCATGATGATTTGACATCAGGGTCCGTTTACGTTG 1080  
Db 1306 GGCTTGTGAAGGGAATCCATGCCATATGATTTGACATCAGGGTCCGTTTACGTTG 1365  
QY 1081 AGGGGCCCAACCTGGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1140  
Db 1366 AGGGGCCCAACCTGGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1425  
QY 1141 GCCCCACCATCTGACACATTCGAGGCTTGGACATACCTGCGGATATGACGCGGAATCC 1200  
Db 1426 GCCCCACCATCTGACACATTCGAGGCTTGGACATACCTGCGGATATGACGCGGAATCC 1485  
QY 1201 ATCTCAAGCTGTGACACGAGCGGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1260  
Db 1486 ATCTCAAGCTGTGACACGAGCGGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1545  
QY 1261 AGGGTCTGGGGACTCTTTCTTGGTGGAGAGGCAAGCTGCTACACAGAGACAAAT 1320  
Db 1546 AGGGTCTGGGGACTCTTTCTTGGTGGAGAGGCAAGCTGCTACACAGAGACAAAT 1605  
QY 1321 GACAAAGTGAACCCAGGAGGAGAACTTTCTGCCAAGTACACAGCTGTGAAGACCTG 1380  
Db 1606 GACAAAGTGAACCCAGGAGGAGAACTTTCTGCCAAGTACACAGCTGTGAAGACCTG 1665  
QY 1381 TGTCAAGCTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440  
Db 1666 TGTCAAGCTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1725  
QY 1441 GAGGAGCCACGGGGAGCTGAAGCTGCATAGTGAAGGGCCCATCGGCTGGGGCGG 1500  
Db 1726 GAGGAGCCACGGGGAGCTGAAGCTGCATAGTGAAGGGCCCATCGGCTGGGGCGG 1785  
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGCAGGCGAGCGCTGACACC 1560  
Db 1786 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGCAGGCGAGCGCTGACACC 1845  
QY 1561 TGTGACAGGGGAGCTAAGCTCAGCTTGGCGGAGCGCGGAGAAAATCTTTCAGGAAG 1620  
Db 1846 TGTGACAGGGGAGCTAAGCTCAGCTTGGCGGAGCGCGGAGAAAATCTTTCAGGAAG 1905  
QY 1621 AAGTACAAGSCCAGTATGTCCGAGTGCCTCATCCGCTCAGTGGCCATCAGGTGGAC 1680  
Db 1906 AAGTACAAGSCCAGTATGTCCGAGTGCCTCATCCGCTCAGTGGCCATCAGGTGGAC 1965  
QY 1681 GGAGGGTGTACACGTAGGCTTGGGTGATGCGGCCAGCCCGAAACCTCACCAAGCGG 1740  
Db 1966 GGAGGGTGTACACGTAGGCTTGGGTGATGCGGCCAGCCCGAAACCTCACCAAGCGG 2025  
QY 1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAAGATGGTGGGACTTTCAGTGGCACT 1800  
Db 2026 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAAGATGGTGGGACTTTCAGTGGCACT 2085  
QY 1801 GGAGGGCTTCCGACTACTCAGCGGCCAACCCCATTTAAAGTGCACATCGGTGCTACATC 1860  
Db 2086 GGAGGGCTTCCGACTACTCAGCGGCCAACCCCATTTAAAGTGCACATCGGTGCTACATC 2145  
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGTACAGCTGTGACCTGAGGCTTGGAAA 1920  
Db 2146 CTAGAGAACGACACAGTCCAGTGTGACCTGTGACCTGTGACCTGAGGCTTGGAAA 2205  
QY 1921 GACCACAAGCTGCACATCGACACGAGATTGAACCTTCAGAAACAAATTAAGAACCTG 1980  
Db 2206 GACCACAAGCTGCACATCGACACGAGATTGAACCTTCAGAAACAAATTAAGAACCTG 2265  
QY 1981 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGGCCAGAAAGATGTGACTGTCACAAAATC 2040  
Db 2266 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGGCCAGAAAGATGTGACTGTCACAAAATC 2325  
QY 2041 AGCTTACACACCCAGCACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGATCTTTTC 2100  
Db 2326 AGCTTACACACCCAGCACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGATCTTTTC 2385

QY 2101 AGGAAGGCGCTGCAAGAGAAAGGACAAAGTGTGCTGTGGCGAGCAGAAAGCAAGAAG 2160  
Db 2386 AGGAAGGCGCTGCAAGAGAAAGGACAAAGTGTGCTGTGGCGAGCAGAAAGCAAGAAG 2445  
QY 2161 AAATCTCGCAAGCTGTCTCAAGCGCTTCAGAAACAAAGACACGTCGACGATGCGGCGCTC 2220  
Db 2446 AAATCTCGCAAGCTGTCTCAAGCGCTTCAGAAACAAAGACACGTCGACGATGCGGCGCTC 2505  
QY 2221 ACGTGCTTCAACCAAGCAACCAAGCACTGGCAGACGCGCGCTTTCTGGACACCTGGGCGCT 2280  
Db 2506 ACGTGTCTTCAACCAAGCAACCAAGCACTGGCAGACGCGCGCTTTCTGGACACCTGGGCGCT 2565  
QY 2281 TTTCTGTGCTGCAACCAAGCAACCAATTAACACGTACTGTGTGCTAGGACCAATCAATGAG 2340  
Db 2566 TTTCTGTGCTGCAACCAAGCAACCAATTAACACGTACTGTGTGCTAGGACCAATCAATGAG 2625  
QY 2341 ACTCACAATTTCTTCTGTGTGAATTTGGAATTTGGCTTCTCTAGAGTACTTTGATCTCAAC 2400  
Db 2626 ACTCACAATTTCTTCTGTGTGAATTTGGAATTTGGCTTCTCTAGAGTACTTTGATCTCAAC 2685  
QY 2401 ACAGACCCCTACCAAGCTGTGAATGCAATGCAATGCACTGACAGGGATGTCTCAACACAG 2460  
Db 2686 ACAGACCCCTACCAAGCTGTGAATGCAATGCAATGCACTGACAGGGATGTCTCAACACAG 2745  
QY 2461 CTACACCTAAGCTGTGAGAGCTGCAAGGCTGCAAGGTTTCAAGCACTGTAAACCCCGG 2520  
Db 2746 CTACACCTAAGCTGTGAGAGCTGCAAGGCTGCAAGGTTTCAAGCACTGTAAACCCCGG 2805  
QY 2521 ACTCGAAACATGAGACCTGGGACTTAAAGATGAGAGGAGCTATGAGCAATACAGGCACTTT 2580  
Db 2806 ACTCGAAACATGAGACCTGGGACTTAAAGATGAGAGGAGCTATGAGCAATACAGGCACTTT 2865  
QY 2581 CAGCGTCCAAAGTGGCCAGAAATGAAAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Db 2866 CAGCGTCCAAAGTGGCCAGAAATGAAAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2925  
QY 2641 GAAGGCTGGGAGGTTAAGAAAACAAAGAGGTGAGCTCCAAAGACATAGAGGACATCAC 2700  
Db 2926 GAAGGCTGGGAGGTTAAGAAAACAAAGAGGTGAGCTCCAAAGACATAGAGGACATCAC 2985  
QY 2701 TGACTGCAAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGAGCCTGTGTCTATTGGCC 2760  
Db 2986 TGACTGCAAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGAGCCTGTGTCTATTGGCC 3045  
QY 2761 AGGAGGCTGAGAAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTGGAGGATTAAC 2820  
Db 3046 AGGAGGCTGAGAAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTGGAGGATTAAC 3105  
QY 2821 CAGCAGGAGCAGAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGATTGA 2880  
Db 3106 CAGCAGGAGCAGAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGATTGA 3165  
QY 2881 TACTCTACAGCTGCAAAATGCAATTTTGTATCAAAAAGTCAACCTAACCCCTCCC 2940  
Db 3166 TACTCTACAGCTGCAAAATGCAATTTTGTATCAAAAAGTCAACCTAACCCCTCCC 3225  
QY 2941 CCAGAGCTCACAAGGAAAACGAGAGAGCGGAGAGAGATTTTCTTGGAAATTTCT 3000  
Db 3226 CCAGAGCTCACAAGGAAAACGAGAGAGCGGAGAGAGATTTTCTTGGAAATTTCT 3285  
QY 3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTCTTAAA 3060  
Db 3286 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTCTTAAA 3345  
QY 3061 TCCTCTTATTTCTTTGTTGTCAAAAGAGAACTTAAGAGCAGGACAGAGGACAGT 3120  
Db 3346 TCCTCTTATTTCTTTGTTGTCAAAAGAGAACTTAAGAGCAGGACAGAGGACAGT 3405  
QY 3121 GGAGAGCTGAAAACAGTGCAGAGACGTTTGAACATGAGTCACTAGCAGCAAAAAGAGATGA 3180  
Db 3406 GGAGAGCTGAAAACAGTGCAGAGACGTTTGAACATGAGTCACTAGCAGCAAAAAGAGATGA 3465  
QY 3181 CATTTACTAGCACTATAAAACCCCTGGTGGCTCTGAAAGAAAACCTGCTTCTATATATAT 3240

Db 3466 CATTACCTAGCACTAATAACCTGTTGCCCTGAGAAATGCGCTTCATTGTATAT 3525  
Qy 3241 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3300  
Db 3526 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3585  
Qy 3301 TTTTTCAGGAGTGTGTGTCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAA 3352  
Db 3586 TTTTTCAGGAGTGTGTGTCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAA 3637

RESULT 6  
ABS56559  
ID ABS56559 standard; cDNA; 4286 BP.

AC ABS56559;  
XX 20-JAN-2003 (first entry)  
XX Full-length human SULF2 cDNA.

Human; SULF2; ss; gene; glucosamine-6-sulphatase; cancer; ischaemia;  
tumour; angiogenesis; coronary; carotid; arterial occlusive disease;  
peripheral arterial disease; atherosclerosis; myointimal hyperplasia;  
thromboangitis obliterans; thrombotic disorder; vasculitis;  
heart attack; myocardial infarction; vascular death; inflammation;  
rheumatoid arthritis; asthma; adult respiratory distress syndrome;  
sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;  
allograft rejection; lymphoma; thrombosis; sulphatase.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 677..3289  
FT CDS /\*tag= a  
FT /product= "Human SULF2 protein"  
XX W0200259327-A2.

XX 01-AUG-2002.

XX 26-DEC-2001; 2001WO-US49793.

XX 27-DEC-2000; 2000US-258577P.

XX 09-FEB-2001; 2001US-267831P.

XX (REGC ) UNIV CALIFORNIA.

XX Rosen SD, Hemmerich S, Tomita M, Palmeri D;

XX WPI; 2002-636541/68.

XX P-PSDB; ABG71836.

XX New sulfatase polypeptides having glucosamine-6-sulfatase activity,  
useful in screening, discovery and preparation of diagnostic and  
therapeutic agents for treating cancer, ischemic conditions,  
inflammation, or thrombosis.

XX Example 4; Fig 10A; 293pp; English.

XX This invention relates to the DNA and protein sequences of a novel  
polypeptide having glucosamine-6-sulphatase activity. The sulphatases of  
are useful in screening, discovery and preparation of diagnostic and  
therapeutic agents for treating cancer, ischaemic conditions,  
inflammation, or thrombosis. The nucleic acids are useful in preparing  
the sulfatase polypeptides, identifying the expression of genes in a  
biological specimen, or generating transgenic non-human animals or  
site-specific gene modification in cell lines. The host cells are  
useful in replicating and/or expressing the polynucleotides or nucleic  
acids. The agents are useful in treating the disorders cited above by  
reducing tumour growth, inflammation, and thrombosis, or increasing  
angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive

CC disease, peripheral arterial disease, atherosclerosis, myointimal  
CC hyperplasia, thromboangitis obliterans, thrombotic disorders,  
CC vasculitis, or preventing ischemic conditions, heart attack (myocardial  
CC infarction), or other vascular death. The sulphatases and/or agents are  
CC also useful in treating rheumatoid arthritis, asthma, adult respiratory  
CC distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple  
CC sclerosis, allograft rejection, and spread of lymphomas to cutaneous  
CC sites. The present sequence represents the full-length cDNA encoding the  
CC mouse SULF2 protein of the invention.

XX  
SQ Sequence 4286 BP; 1188 A; 1145 C; 1108 G; 845 T; 0 other;

Query Match 99.5%; Score 3346.8; DB 24; Length 4286;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GACTCCCGCATCCCAAGAGAGCAGCAGATCAGCAAAAAGAGATGGGCCCGGAGC 60  
Db 632 GACTCCCGCATCCCAAGAGAGCAGCAGATCAGCAAAAAGAGATGGGCCCGGAGC 691  
Qy 61 CTGCTGCTGTGCTTGTCTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTC 120  
Db 692 CTGCTGCTGTGCTTGTCTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTC 751  
Qy 121 CTGTGCGCACACCGCTGAAAGCAGGTTTCAGAGGACCGCAGGACATCCGCCCAAC 180  
Db 752 CTGTGCGCACACCGCTGAAAGCAGGTTTCAGAGGACCGCAGGACATCCGCCCAAC 811  
Qy 181 ATCATCTGTGTGTGACGAGCAGCAGGATGTGGAGCTGGGTTCATGCAAGGTGATGAAC 240  
Db 812 ATCATCTGTGTGTGACGAGCAGCAGGATGTGGAGCTGGGTTCATGCAAGGTGATGAAC 871  
Qy 241 AAGACCCGGCGCATATGAGAGCGAGCGGGCGGCGCATTTATCAACGGCTTCGTGACCA 300  
Db 872 AAGACCCGGCGCATATGAGAGCGAGCGGGCGGCGCATTTATCAACGGCTTCGTGACCA 931  
Qy 301 CCCATGTGTGCCCTTCAACGCTCTCCATCTCTCACTGGGAAGTACGTCCACACCAAC 360  
Db 932 CCCATGTGTGCCCTTCAACGCTCTCCATCTCTCAACGGGAAGTACGTCCACACCAAC 991  
Qy 361 ACCTACACCAACAAATGAGAACTGTCTCTCGCCCTCTCTGGCAGCAGCAGAGAGCGCG 420  
Db 992 ACCTACACCAACAAATGAGAACTGTCTCTCGCCCTCTCTGGCAGCAGCAGAGAGCGCG 1051  
Qy 421 ACCTTTCGGGTGTACCTCAATAGCAGTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480  
Db 1052 ACCTTTCGGGTGTACCTCAATAGCAGTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 1111  
Qy 481 AATGAATACAAACGGCTCTCTACGTGCCACCCGGCTGGGAAGAGTGGGTTCGACTCTTAAA 540  
Db 1112 AATGAATACAAACGGCTCTCTACGTGCCACCCGGCTGGGAAGAGTGGGTTCGACTCTTAAA 1171  
Qy 541 AACTCCCGCTTTTATACTACAGCTGTGTGGAAAGGGGTGAAAGAAAGCAGCGCTCC 600  
Db 1172 AACTCCCGCTTTTATACTACAGCTGTGTGGAAAGGGGTGAAAGAAAGCAGCGCTCC 1231  
Qy 601 GACTACTCCCAAGGATTACCTCACAGACCTCATACCAATGACAGCGGTGAGCTTCTTCGCG 660  
Db 1232 GACTACTCCCAAGGATTACCTCACAGACCTCATACCAATGACAGCGGTGAGCTTCTTCGCG 1291  
Qy 661 ACGTCCCAAGAGATGTACCCGCAAGCCAGTCTCATGTGTATCAGCCATGAGCGCCCC 720  
Db 1292 ACGTCCCAAGAGATGTACCCGCAAGCCAGTCTCATGTGTATCAGCCATGAGCGCCCC 1351  
Qy 721 CACGGCCCTGAGGATTACGCCCCCAATATTACGCGCTCTTCCCAACGCAATCTTCAGCAC 780  
Db 1352 CACGGCCCTGAGGATTACGCCCCCAATATTACGCGCTCTTCCCAACGCAATCTTCAGCAC 1411  
Qy 781 ATCAGCGCGAGCTCAACTAGCGGCCCAACCCGGAACAAACACTGGATCATGCGGTACAG 840  
Db 1412 ATCAGCGCGAGCTCAACTAGCGGCCCAACCCGGAACAAACACTGGATCATGCGGTACAG 1471  
Qy 841 GGGCCCATGAAGCCCATCCATGGAATTTCACCAACATGCTCCAGCGGAAGCGGCTTCAG 900

1472 GGGCCCATGAAGCCATCCATCGAATTTCAACAATGCTCCAGCGAAGCGCTTGCGAG 1531  
901 ACCCTCATGTCGGTGGAGCACTCCATGGAGAGATTTACATCATGCTGGTTGAGACGGC 960  
1532 ACCCTCATGTCGGTGGAGCACTCCATGGAGAGATTTACATCATGCTGGTTGAGACGGC 1591  
961 GAGCTGGACAACACGTACATCGTATACACCGCGGACCGGTACACATCGGCGCAGTTT 1020  
1592 GAGCTGGACAACACGTACATCGTATACACCGCGGACCGGTACACATCGGCGCAGTTT 1651  
1021 GGCCTGGTGAAGGGAATCCATGCGATAGTTGATGATGAGGTCCCGTTCTACGTG 1080  
1652 GGCCTGGTGAAGGGAATCCATGCGATAGTTGATGATGAGGTCCCGTTCTACGTG 1711  
1081 AGGGGCCCCAAGCTGGAAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1140  
1712 AGGGGCCCCAAGCTGGAAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1771  
1141 GCGCCACCATCTGTGACATTTGAGGCTGTGACATACCTGCGGATATGGACGGGAATCC 1200  
1772 GCGCCACCATCTGTGACATTTGAGGCTGTGACATACCTGCGGATATGGACGGGAATCC 1831  
1201 ATCTCAAGCTGCTGGACACGAGCGGCGGCTGAATCGTTTCACTTCAAAAAGAGATG 1260  
1832 ATCTCAAGCTGCTGGACACGAGCGGCGGCTGAATCGTTTCACTTCAAAAAGAGATG 1891  
1261 AGGGTCTGGCGGATCTCTTCTTGTGGAGAGCAAGCTGTCTACACAGAGAGACAAT 1320  
1892 AGGGTCTGGCGGATCTCTTCTTGTGGAGAGCAAGCTGTCTACACAGAGAGACAAT 1951  
1321 GACAAAGTGGAGCGCCAGAGAGAGAACTTCTGCGCCAAAGTACCAAGCGTGTCAAGGACCTG 1380  
1952 GACAAAGTGGAGCGCCAGAGAGAGAACTTCTGCGCCAAAGTACCAAGCGTGTCAAGGACCTG 2011  
1381 TGTACAGCTGTGAGTACACAGCGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440  
2012 TGTACAGCTGTGAGTACACAGCGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2071  
1441 GAGGACGCCAGCGGAGAGCTGAAGCTGATAGTGCAGAGGCGCCCATCGGCTGGCGGC 1500  
2072 GAGGACGCCAGCGGAGAGCTGAAGCTGATAGTGCAGAGGCGCCCATCGGCTGGCGGC 2131  
1501 AGCAGAGCCCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCAACC 1560  
2132 AGCAGAGCCCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCAACC 2191  
1561 TGTGACAGCGGGAATAAGCTCAGCTGGCGGACCGCCGGAAGAACTCTTCAAGAAG 1620  
2192 TGTGACAGCGGGAATAAGCTCAGCTGGCGGACCGCCGGAAGAACTCTTCAAGAAG 2251  
1621 AAGTACAGGCGCAGTATGTCGCGAGTCTGCTCCATCGCTCAGTGGCCATCGAGGTGAC 1680  
2252 AAGTACAGGCGCAGTATGTCGCGAGTCTGCTCCATCGCTCAGTGGCCATCGAGGTGAC 2311  
1681 GGCAGGGTGTACCAAGCTGAGGCTGATGCGCCAGGCGCCGAAACCTTCAAGAGCGG 1740  
2312 GGCAGGGTGTACCAAGCTGAGGCTGATGCGCCAGGCGCCGAAACCTTCAAGAGCGG 2371  
1741 CACTGGCAGGGCCCTGAGGACCAAGATGACAGATGATGGGACTTCAGTGGCACT 1800  
2372 CACTGGCAGGGCCCTGAGGACCAAGATGACAGATGATGGGACTTCAGTGGCACT 2431  
1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAGTGAACATCGGTGTACATC 1860  
2432 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAGTGAACATCGGTGTACATC 2491  
1861 CTAGAGACGACACAGTCCAGTGTGACTGACCTGTGTAAGTCCCTGAGGCTTGAAA 1920  
2492 CTAGAGACGACACAGTCCAGTGTGACTGACCTGTGTAAGTCCCTGAGGCTTGAAA 2551  
1921 GACCAAGCTGCACATCGACCAAGATTTGAACCCCTGCGAGACAAATTAAGAACCTG 1980

2552 GACCAAGCTGCACATCGACACAGAGATTGAAACCCCTGCAGAAACAAATTAAGAACCTG 2611  
1981 AGGAGTCCGAGGTCACTGAGAAAGAGCGCCAGAGAAATGCTGCTGCAAAATC 2040  
2612 AGGAGTCCGAGGTCACTGAGAAAGAGCGCCAGAGAAATGCTGCTGCAAAATC 2671  
2041 AGTATCACACACCCAGCACAAAGGCGCCCTCAAGCACAGAGGCTTCCAGTCTGATCCTTTC 2100  
2672 AGTATCACACACCCAGCACAAAGGCGCCCTCAAGCACAGAGGCTTCCAGTCTGATCCTTTC 2731  
2101 AGGAGGCTCCAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAGAGAG 2160  
2732 AGGAGGCTCCAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAGAGAG 2791  
2161 AAATCTCCGAGCTGTCAAGGCGCTCGAGAACCAACGACACGTCAGCATGCCAGGCTC 2220  
2792 AAATCTCCGAGCTGTCAAGGCGCTCGAGAACCAACGACACGTCAGCATGCCAGGCTC 2851  
2221 AGTGTCTTCCACACACACACGACACGACACGCGCTTCTGGACACCTGGGCGCT 2280  
2852 AGTGTCTTCCACACACACACGACACGACACGCGCTTCTGGACACCTGGGCGCT 2911  
2281 TTCTGTGCTTGCACACGCGCCCAACATTAACGCTACTGTGTCATGAGGACCATCAATGAG 2340  
2912 TTCTGTGCTTGCACACGCGCCCAACATTAACGCTACTGTGTCATGAGGACCATCAATGAG 2971  
2341 ACTCAAAATTTCTTCTGTGTGAATTTGCAATGCGCTTCTTAGAGTACTTGTGATCTCAAC 2400  
2972 ACTCAAAATTTCTTCTGTGTGAATTTGCAATGCGCTTCTTAGAGTACTTGTGATCTCAAC 3031  
2401 ACAGACCCCTTACAGCTGATGATGAGTGAACACACTGGAAGGATGCTCTCAACACAG 2460  
3032 ACAGACCCCTTACAGCTGATGATGAGTGAACACACTGGAAGGATGCTCTCAACACAG 3091  
2461 CTACAGCTCAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520  
3092 CTACAGCTCAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 3151  
2521 ACTCGAAAATGAGACCTTGAAGTGAAGGAGCTATGAGCAATACAGGCGATTT 2580  
3152 ACTCGAAAATGAGACCTTGAAGTGAAGGAGCTATGAGCAATACAGGCGATTT 3211  
2581 CAGCGTGGAAAGTGGCGAGAAATGAAGAGACCTTCTTCAAAATCACTGGGACCACTGTCG 2640  
3212 CAGCGTGGAAAGTGGCGAGAAATGAAGAGACCTTCTTCAAAATCACTGGGACCACTGTCG 3271  
2641 GAAGGCTGGGAGGTTAAGAAACAAAGAGGTGGACCTTCAAAAACATAGAGGCACTCAC 2700  
3272 GAAGGCTGGGAGGTTAAGAAACAAAGAGGTGGACCTTCAAAAACATAGAGGCACTCAC 3331  
2701 TGACTGCACAGCAATGAAGAACCAATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760  
3332 TGACTGCACAGCAATGAAGAACCAATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 3391  
2761 AGGAGGCTTGAGAAAGCAAGCACTCTCAGTCAACATGACAGATTTGAGGAGATAAC 2820  
3392 AGGAGGCTTGAGAAAGCAAGCACTCTCAGTCAACATGACAGATTTGAGGAGATAAC 3451  
2821 CAGCAGGACAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
3452 CAGCAGGACAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3511  
2881 TACTCACCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTAACCTCCC 2940  
3512 TACTCACCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTAACCTCCC 3571  
2941 CAGAGAGCTCACAAGAGGAAAACGGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000  
3572 CAGAGAGCTCACAAGAGGAAAACGGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3631  
3001 CCCAAGGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAA 3060  
3632 CCCAAGGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAA 3691



QY 3061 TCCTCTATTCTTTTGGTTTGTACAAAGAGGAAGAACTAAGAGGAGGACAGAGGCAACGT 3120  
Db 3692 TCCTCTATTCTTTTGGTTTGTACAAAGAGGAAGAACTAAGAGGAGGACAGAGGCAACGT 3751  
QY 3121 GGAGAGGCTGAAAACAGTGCGAGAGAGCTTTGACAAATGAGTCAGTACGACAAAGAGATGA 3180  
Db 3752 GGAGAGGCTGAAAACAGTGCGAGAGAGCTTTGACAAATGAGTCAGTACGACAAAGAGATGA 3811  
QY 3181 CATTACCTAGCACTATAACCCCTGGTTGCTCTGAAGAACTGCTTCATTGTATATAT 3240  
Db 3812 CATTACCTAGCA-TATAAACCCCTGGTTGCTCTGAAGAACTGCTTCATTGTATATAT 3870  
QY 3241 GTGACTATTTTACATGTAATCAACATCGGAACCTTTTAGGGGAACCTTAATAGAAATCCCAA 3300  
Db 3871 GTGACTATTTTACATGTAATCAACATCGGAACCTTTTAGGGGAACCTTAATAGAAATCCCAA 3930  
QY 3301 TTTTCAGGAGTGCTGTAATAAAGCTCTGTGGCCAGTGTAAAGAAAGAAAGAAA 3360  
Db 3931 TTTTCAGGAGTGCTGTAATAAAGCTCTGTGGCCAGTGTAAAGAAAGAAAGAAA 3990  
QY 3361 AA 3362  
Db 3991 AA 3992

RESULT 7  
AAH76201  
ID AAH76201 standard; cDNA; 3348 BP.  
AC AAH76201;  
XX 29-OCT-2001 (first entry)  
XX Human drug metabolizing enzyme encoding cDNA (ID No. 1559210CB1).  
DE Drug metabolizing enzyme; DME; immunosuppressive; cytostatic; ophthalmic;  
KW hepatocytrophic; antiallergic; antiaesthetic; antibacterial; antiviral;  
KW antisense therapy; gene therapy; human; ss.  
XX Homo sapiens.

Key Location/Qualifiers  
CDS 35..2647  
FT /\*tag= a  
XX WO200159127-A2.  
XX 16-AUG-2001.  
XX 08-FEB-2001; 2001WO-US04423.  
XX 11-FEB-2000; 2000US-0181856.  
XX 17-FEB-2000; 2000US-0183684.  
XX 25-FEB-2000; 2000US-0185141.  
XX 03-MAR-2000; 2000US-0186818.  
XX 09-MAR-2000; 2000US-0188345.  
XX 17-MAR-2000; 2000US-0189997.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lal P;  
XX Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JL;  
XX WPI; 2001-514673/56.  
XX P-PSDB; AAB85774.  
XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the  
XX diagnosis, treatment, and prevention of autoimmune/inflammatory, cell  
XX proliferative, developmental and endocrine disorders -  
XX Claim 5; Page 144-145; 150pp; English.

CC The invention provides human drug metabolizing enzymes (DME) and  
CC polynucleotides encoding the DMEs. The DME can be expressed by standard  
CC recombinant methodology. DMEs and their agonists and antagonists are  
CC useful for the diagnosis, treatment, and prevention of autoimmune/  
CC inflammatory, cell proliferative, developmental, endocrine such as  
CC aneurysm, eye, metabolic, and gastrointestinal disorders, including liver  
CC disorders and infection. The present sequence represents a human DME  
CC encoding cDNA.

XX  
SQ Sequence 3348 BP; 927 A; 933 C; 869 G; 619 T; 0 other;  
Query Match 99.58; Score 3344.8; DB 22; Length 3348;  
Best Local Similarity 99.94; Pred. No. 0;  
Matches 3346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 12 CCCAAAGAGAGCAGATCAGCAAAAGAGATGGGCCCCCGAGGCTCGTGTGTG 71  
Db 1 CCCAAAGAGAGCAGATCAGCAAAAGAGATGGGCCCCCGAGGCTCGTGTGTG 60  
QY 72 CTTCCTGTCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTCTGTGCAACA 131  
Db 61 CTTCCTGTCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTCTGTGCAACA 120  
QY 132 CCGCTCAAGAGGAGGTTTCAGAGGAGCCGCAAGAACATCCGCCCAACATCATCTCTGT 191  
Db 121 CCGCTCAAGAGGAGGTTTCAGAGGAGCCGCAAGAACATCCGCCCAACATCATCTCTGT 180  
QY 192 GCTGACGAGCAGCAGGATGTGAGCTGGTTCATGCAAGTGTGATGAAACCGGCG 251  
Db 181 GCTGACGAGCAGCAGGATGTGAGCTGGTTCATGCAAGTGTGATGAAACCGGCG 240  
QY 252 CATCATGAGCAGGCGGGGCGGCACTTTCATCAAGCTTCGTGACCAACCCATGTCTG 311  
Db 241 CATCATGAGCAGGCGGGGCGGCACTTTCATCAAGCTTCGTGACCAACCCATGTCTG 300  
QY 312 CCCCTCAGCTCTCTCCATCTCTGCAAGTACGTCCAAACCAACCAACCAACCAACAA 371  
Db 301 CCCCTCAGCTCTCTCCATCTCTGCAAGTACGTCCAAACCAACCAACCAACCAACAA 360  
QY 372 CAATGAGACTGCTCTCCGCTCTCTGCGAGGAGCAGCAGGAGCGGCGGCTTTCGCT 431  
Db 361 CAATGAGACTGCTCTCCGCTCTCTGCGAGGAGCAGCAGGAGCGGCGGCTTTCGCT 420  
QY 432 GTACCTCAATAGCACTGGCTACCGGAGAGCTTTCTTCGGGAAGTATCTTAATGAATCAA 491  
Db 421 GTACCTCAATAGCACTGGCTACCGGAGAGCTTTCTTCGGGAAGTATCTTAATGAATCAA 480  
QY 492 CGGCTCTTACGTCGCAACCGGCTGGAAGAGTGGTGGAGCTCTTAAACACTCCGCTT 551  
Db 481 CGGCTCTTACGTCGCAACCGGCTGGAAGAGTGGTGGAGCTCTTAAACACTCCGCTT 540  
QY 552 TTATAACTACACGCTGTGCGAAGCGGCTGAAAGAAAGCAGGCTCCGACTACTCCAA 611  
Db 541 TTATAACTACACGCTGTGCGAAGCGGCTGAAAGAAAGCAGGCTCCGACTACTCCAA 600  
QY 612 GGATTAACCTCAACAGCTCTATCAACAAATGACAGCTGAGTCTTTCGCAAGTCAAGAA 671  
Db 601 GGATTAACCTCAACAGCTCTATCAACAAATGACAGCTGAGTCTTTCGCAAGTCAAGAA 660  
QY 672 GATGTACCCGACAGGCGGCTCTCATGCTCATGAGCCATGCGAGGCGGCGGCTTGA 731  
Db 661 GATGTACCCGACAGGCGGCTCTCATGCTCATGAGCCATGCGAGGCGGCGGCTTGA 720  
QY 732 GGATTCAGCCCCCAATATTCACGCTCTTCCCAAGCATCTCAGACATCAGCGGAG 791  
Db 721 GGATTCAGCCCCCAATATTCACGCTCTTCCCAAGCATCTCAGACATCAGCGGAG 780  
QY 792 CTACAACTACGCGGCGGCGGCGGCAACACTGATCATGGCTTACAGGGGCGGCGGCGG 851  
Db 781 CTACAACTACGCGGCGGCGGCGGCAACACTGATCATGGCTTACAGGGGCGGCGGCGG 840  
QY 852 GCCCATCCACATGGAATTCACCAACATGCTCTCAGCGGAAGCGCTTGAGAGCCCTCATGTC 911



Db 841 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGAAGCGCTTGAGACACCTCATGTC 900  
Qy 912 GGTGAGAGACTCCATGGAGAGAGATTTCACCATGCTGTTTCAGACGCGGCGAGCTGGACAA 971  
Db 901 GGTGAGAGACTCCATGGAGAGAGATTTCACCATGCTGTTTCAGACGCGGCGAGCTGGACAA 960  
Qy 972 CACGTACATCGTATACACCGCGGACCAACGGTTACACATCGCCAGTTTGGCCCTGGTGA 1031  
Db 961 CACGTACATCGTATACACCGCGGACCAACGGTTACACATCGCCAGTTTGGCCCTGGTGA 1020  
Qy 1032 AGGGAATTCATGCCATATGATTTGACATCAGGGTCCCGTTTCTACGTGAGGGGCCCA 1091  
Db 1021 AGGGAATTCATGCCATATGATTTGACATCAGGGTCCCGTTTCTACGTGAGGGGCCCA 1080  
Qy 1092 CGTGAAGCCGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1151  
Db 1081 CGTGAAGCCGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1140  
Qy 1152 CTTGAGACATTCGAGCCCTGGACATACCTGCGGATATGACCGGGAATCCATCTCAAGCT 1211  
Db 1141 CTTGAGACATTCGAGCCCTGGACATACCTGCGGATATGACCGGGAATCCATCTCAAGCT 1200  
Qy 1212 GCTGACACGAGCGCGCTGAAATCGGTTTCACTTGAAAGAAAGATGAGGGTCTGGCG 1271  
Db 1201 GCTGACACGAGCGCGCTGAAATCGGTTTCACTTGAAAGAAAGATGAGGGTCTGGCG 1260  
Qy 1272 GGAATCTCTTGTGTGAGAGAGAGCAAGCTGCTACACAGAGAGACAATGACAAAGTGA 1331  
Db 1261 GGAATCTCTTGTGTGAGAGAGAGCAAGCTGCTACACAGAGAGACAATGACAAAGTGA 1320  
Qy 1332 CGCCAGAGAGAGAACTTCTGCCCCAAGTACACAGCTGTGAGAGACCTGTGTCAGCGTGC 1391  
Db 1321 CGCCAGAGAGAGAACTTCTGCCCCAAGTACACAGCTGTGAGAGACCTGTGTCAGCGTGC 1380  
Qy 1392 TGAGTACACAGAGCGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1451  
Db 1381 TGAGTACACAGAGCGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Qy 1452 GGGAGAGTGAAGCTGCATAGTGCAGAGGCGCCATGCGGCTGGCGCGGAGCAGAGCCCT 1511  
Db 1441 GGGAGAGTGAAGCTGCATAGTGCAGAGGCGCCATGCGGCTGGCGCGGAGCAGAGCCCT 1500  
Qy 1512 CTCCAACCTCGTGCCTCCAGTACTACGGGCGAGGCGAGGCGCTGCACTGTGACAGCGG 1571  
Db 1501 CTCCAACCTCGTGCCTCCAGTACTACGGGCGAGGCGAGGCGCTGCACTGTGACAGCGG 1560  
Qy 1572 GGAATCAAGCTCAGCTGCGCGGAGAGCGCGGAAACCTTCAAGAGAGAGTACAAGGC 1631  
Db 1561 GGAATCAAGCTCAGCTGCGCGGAGAGCGCGGAAACCTTCAAGAGAGAGTACAAGGC 1620  
Qy 1632 CAGCTATGTCGCGAGTCCATCGCTCAGTGGCCATCGAGGTGACGCGAGGCTGTA 1691  
Db 1621 CAGCTATGTCGCGAGTCCATCGCTCAGTGGCCATCGAGGTGACGCGAGGCTGTA 1680  
Qy 1692 CCACTAGGCTGGGTGATGCGCCGAGCGGCGGAAACCTTCAAGAGAGAGTACAAGGC 1751  
Db 1681 CCACTAGGCTGGGTGATGCGCCGAGCGGCGGAAACCTTCAAGAGAGAGTACAAGGC 1740  
Qy 1752 GGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCACTGCGCACTGGAGGCTTCC 1811  
Db 1741 GGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCACTGCGCACTGGAGGCTTCC 1800  
Qy 1812 CGACTACTCAGCGGCAACCCCATTAAGTGAACATCGTGTGCTACATCTTAGAGACGA 1871  
Db 1801 CGACTACTCAGCGGCAACCCCATTAAGTGAACATCGTGTGCTACATCTTAGAGACGA 1860  
Qy 1872 CACAGTCCAGTGTGACCTGGAAGCTGTACAGTCCCTGAGGCGCTGGAGACCAAGCT 1931  
Db 1861 CACAGTCCAGTGTGACCTGGAAGCTGTACAGTCCCTGAGGCGCTGGAGACCAAGCT 1920  
Qy 1932 GCACATCCAGCAGAGATTGAACCTCGAGAACCAAAATTAAGAACCTTAGGGAGTCCG 1991  
Db 1921 GCACATCCAGCAGAGATTGAACCTCGAGAACCAAAATTAAGAACCTTAGGGAGTCCG 1980

Qy 1992 AGGTCACTTGAAGAAAAAGCGCCAGAGAAATGTGACTGTCAAAAAATCAGCTACACAC 2051  
Db 1981 AGGTCACTTGAAGAAAAAGCGCCAGAGAAATGTGACTGTCAAAAAATCAGCTACACAC 2040  
Qy 2052 CCAGCAAAAAGCCCGCTCAAGCAAGAGGCTCCAGTCTGCAATCTTTTCAGGAAGGGCT 2111  
Db 2041 CCAGCAAAAAGCCCGCTCAAGCAAGAGGCTCCAGTCTGCAATCTTTTCAGGAAGGGCT 2100  
Qy 2112 GCAAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAAGAGAAATCAGCA 2171  
Db 2101 GCAAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAAGAGAAATCAGCA 2160  
Qy 2172 GTGTCTCAAGCGCTTCAGAAACAACGACACGTCGACGATGCGAGGCTCACGTGTTTCA 2231  
Db 2161 GTGTCTCAAGCGCTTCAGAAACAACGACACGTCGACGATGCGAGGCTCACGTGTTTCA 2220  
Qy 2232 CCAAGCAACAGCACTGGCAGACGGCGCTTCTTGGACACTGGGGCTTCTGTGCTG 2291  
Db 2221 CCAAGCAACAGCACTGGCAGACGGCGCTTCTTGGACACTGGGGCTTCTGTGCTG 2280  
Qy 2292 CACAGCGCAACAATAACAACGTACTGTGTCATGAGGACCAATCAATGAGACTCACAAAT 2351  
Db 2281 CACAGCGCAACAATAACAACGTACTGTGTCATGAGGACCAATCAATGAGACTCACAAAT 2340  
Qy 2352 CTTCTTCTGTGAATTTGMACTGGCTTCTTAGAGTACTTTGATCTCAACAGACCCCTA 2411  
Db 2341 CTTCTTCTGTGAATTTGMACTGGCTTCTTAGAGTACTTTGATCTCAACAGACCCCTA 2400  
Qy 2412 CAGCTGATGAATGACGTGAACAACACTGGACAGGAGTGTCTCAACAGCTACAGTACA 2471  
Db 2401 CAGCTGATGAATGACGTGAACAACACTGGACAGGAGTGTCTCAACAGCTACAGTACA 2460  
Qy 2472 GTCATGAGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGAGCTCGAAA 2531  
Db 2461 GTCATGAGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGAGCTCGAAA 2520  
Qy 2532 GGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCAAGTTTACGCTCGAAA 2591  
Db 2521 GGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCAAGTTTACGCTCGAAA 2580  
Qy 2592 GTGGCCAGAAATGAAGAGACCTTCTTCAAAATCACTGGGACAACTGTGGGAAGGCTGGA 2651  
Db 2581 GTGGCCAGAAATGAAGAGACCTTCTTCAAAATCACTGGGACAACTGTGGGAAGGCTGGA 2640  
Qy 2652 AGGTTAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACTGACTGCACAG 2711  
Db 2641 AGGTTAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACTGACTGCACAG 2700  
Qy 2712 GCAATGAAAAACCATGTGGTGAATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGA 2771  
Db 2701 GCAATGAAAAACCATGTGGTGAATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGA 2760  
Qy 2772 GAAAGCAGCAGCCTCTCAGTCAACATGACAGATTCCTGAGGATTAACAGCAGAGCA 2831  
Db 2761 GAAAGCAGCAGCCTCTCAGTCAACATGACAGATTCCTGAGGATTAACAGCAGAGCA 2820  
Qy 2832 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTTATACCTCACAG 2891  
Db 2821 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTTATACCTCACAG 2880  
Qy 2892 CTGCAAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTAACCTCCCGCAGAGGCTCA 2951  
Db 2881 CTGCAAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTAACCTCCCGCAGAGGCTCA 2940  
Qy 2952 CAAAGGAAAACGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGCA 3011  
Db 2941 CAAAGGAAAACGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGCA 3000  
Qy 3012 AAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAATCTCTTATTTC 3071  
Db 3001 AAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAATCTCTTATTTC 3060

QY 3072 TTTTGGTTTGTCTCAAGAAGAGAACTAAGAAGCAGGACAGAGCAACGCTGGAGAGCTTGA 3131  
Db 3061 TTTTGGTTTGTCTCAAGAAGAGAACTAAGAAGCAGGACAGAGCAACGCTGGAGAGCTTGA 3120  
QY 3132 AAACAGTGCAGAGACGTTTGACAAATGAGTCAGTAGCAGCAAAAAGAGATGACATTTACCTAG 3191  
Db 3121 AAACAGTGCAGAGACGTTTGACAAATGAGTCAGTAGCAGCAAAAAGAGATGACATTTACCTAG 3180  
QY 3192 CACTATAAACCCCTGGTTGCTCTGGAAGAACTGCTTCAATGATATATATGACTATTTA 3251  
Db 3181 CACTATAAACCCCTGGTTGCTCTGGAAGAACTGCTTCAATGATATATATGACTATTTA 3240  
QY 3252 CATGTAATCAACATGGGAACCTTTTAGGGGAACCTTAATAAGAAATCCCAATTTTCAGGAGT 3311  
Db 3241 CATGTAATCAACATGGGAACCTTTTAGGGGAACCTTAATAAGAAATCCCAATTTTCAGGAGT 3300  
QY 3312 GGTGGTGTCAATAAAGCGCTCTGGCCAGTCAGTAAAGAAAAA 3359  
Db 3301 GGTGGTGTCAATAAAGCGCTCTGGCCAGTCAGTAAAGAAAAA 3348

## RESULT 8

AAZ64969

ID AAZ64969 standard; cDNA; 3906 BP.

XX AC AAZ64969;

XX XX

DT 05-APR-2000 (first entry)

XX XX

DE Membrane-bound protein Prol120 encoding cDNA.

XX XX

KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

KW KW

XX pharmaceutical; receptor immunoadhesin; gene mapping; ss.

XX XX

OS Homo sapiens.

XX XX

PN WO9963088-A2.

XX XX

PD 09-DEC-1999.

XX XX

PF 02-JUN-1999; 99WO-US12252.

XX XX

XX 02-JUN-1998; 98US-0087607.

PR PR

PR 02-JUN-1998; 98US-0087609.

PR PR

PR 03-JUN-1998; 98US-0087759.

PR PR

PR 04-JUN-1998; 98US-0087827.

PR PR

PR 04-JUN-1998; 98US-0088021.

PR PR

PR 04-JUN-1998; 98US-0088025.

PR PR

PR 04-JUN-1998; 98US-0088028.

PR PR

PR 04-JUN-1998; 98US-0088030.

PR PR

PR 04-JUN-1998; 98US-0088033.

PR PR

PR 04-JUN-1998; 98US-0088326.

PR PR

PR 05-JUN-1998; 98US-0088467.

PR PR

PR 05-JUN-1998; 98US-0088202.

PR PR

PR 05-JUN-1998; 98US-0088212.

PR PR

PR 05-JUN-1998; 98US-0088217.

PR PR

PR 09-JUN-1998; 98US-0088655.

PR PR

PR 10-JUN-1998; 98US-0088722.

PR PR

PR 10-JUN-1998; 98US-0088730.

PR PR

PR 10-JUN-1998; 98US-0088734.

PR PR

PR 10-JUN-1998; 98US-0088738.

PR PR

PR 10-JUN-1998; 98US-0088740.

PR PR

PR 10-JUN-1998; 98US-0088741.

PR PR

PR 10-JUN-1998; 98US-0088742.

PR PR

PR 10-JUN-1998; 98US-0088810.

PR PR

PR 10-JUN-1998; 98US-0088811.

PR PR

PR 10-JUN-1998; 98US-0088824.

PR PR

PR 10-JUN-1998; 98US-0088825.

PR PR

PR 10-JUN-1998; 98US-0088826.

PR PR

PR 11-JUN-1998; 98US-0088858.

PR PR

PR 11-JUN-1998; 98US-0088861.

PR PR

PR 11-JUN-1998; 98US-0088863.

PR PR

PR 11-JUN-1998; 98US-0088876.  
PR 12-JUN-1998; 98US-0089090.  
PR 12-JUN-1998; 98US-0089105.  
PR 16-JUN-1998; 98US-0089440.  
PR 16-JUN-1998; 98US-0089512.  
PR 16-JUN-1998; 98US-0089514.  
PR 17-JUN-1998; 98US-0089532.  
PR 17-JUN-1998; 98US-0089538.  
PR 17-JUN-1998; 98US-0089598.  
PR 17-JUN-1998; 98US-0089599.  
PR 17-JUN-1998; 98US-0089600.  
PR 17-JUN-1998; 98US-0089653.  
PR 18-JUN-1998; 98US-0089801.  
PR 18-JUN-1998; 98US-0089907.  
PR 18-JUN-1998; 98US-0089908.  
PR 19-JUN-1998; 98US-0089947.  
PR 19-JUN-1998; 98US-0089948.  
PR 19-JUN-1998; 98US-0089952.  
PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 25-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 01-JUL-1998; 98US-0091544.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.











QY 3001 CCCAAGGCGGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3060  
Db 3554 CCCAAGGCGGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3613  
QY 3061 TCCTCTTATTTCTTTGGTTGTCACAAAGAGGAAGTAAAGACAGACAGAGGCAACGT 3120  
Db 3614 TCCTCTTATTTCTTTGGTTGTCACAAAGAGGAAGTAAAGACAGACAGAGGCAACGT 3673  
QY 3121 GGAGAGGCTGAAAACAGTCGACAGAGAGTTCACAAATGAGTCAGTAGCACAAAAGAGATGA 3180  
Db 3674 GGAGAGGCTGAAAACAGTCGACAGAGAGTTCACAAATGAGTCAGTAGCACAAAAGAGATGA 3733  
QY 3181 CATTACCTAGACATATAAACCCCTGGTTGCTCTGAAAGAACTGCCCTTCATGATATAT 3240  
Db 3734 CATTACCTAGACATATAAACCCCTGGTTGCTCTGAAAGAACTGCCCTTCATGATATAT 3793  
QY 3241 GTGACTATTATCATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3300  
Db 3794 GTGACTATTATCATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3853  
QY 3301 TTTTCAGAGTGGTGGTGCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3353  
Db 3854 TTTTCAGAGTGGTGGTGCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3906  
RESULT 10  
ID AAF44115  
XX AAF44115 standard; cDNA; 3906 BP.  
AC AAF44115;  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE Human PRO1120 (UNQ559) nucleotide sequence SEQ ID NO:83.  
XX  
KW Human; secreted and transmembrane protein; PRO; cytostatic;  
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
KW diagnostic assay; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200073454-A1.  
XX  
PD 07-DEC-2000.  
XX  
PF 30-MAR-2000; 2000WO-US08439.  
XX  
PR 02-JUN-1999; 99WO-US12252.  
PR 23-JUN-1999; 99US-0141037.  
PR 07-JUL-1999; 99US-0143048.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 28-JUL-1999; 99US-0146222.  
PR 17-AUG-1999; 99US-0149396.  
PR 15-SEP-1999; 99WO-US21090.  
PR 08-OCT-1999; 99US-0158663.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 11-FEB-2000; 2000WO-US00376.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 02-MAR-2000; 2000WO-US05004.  
PR 15-MAR-2000; 2000WO-US05841.  
PR 20-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
XX

(GETH ) GENENTECH INC.  
Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
Grimaldi CJ, Gurney AL, Kljavin IC, Raptier MA, Pan J, Paoni NF;  
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
Zhang Z;  
WPI; 2001-032160/04.  
P-PSDB; AAB65171.  
PRO polynucleotides used to produce polypeptides used to target  
bioactive molecules such as toxins, radiolabels or antibodies, to  
specific cells, to cause targeted cell death -  
Claim 2; Fig 46; 935pp; English.  
The present invention describes human secreted and transmembrane PRO  
proteins. The PRO proteins have cytostatic activity. The PRO proteins  
can be used for targeted delivery of bioactive molecules, such as  
toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
sequences, and their fragments, can be used as hybridisation probes, in  
chromosomal and gene mapping, and in the generation of anti-sense RNA  
and DNA. They may also be used to produce transgenic animals which are  
used to develop and screen therapeutically useful reagents. The PRO  
nucleotide and protein sequence can be used for tissue typing and in  
treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
AAF44270 to AAF44470 represent PCR primers and hybridisation probes used  
in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein  
sequences given in the exemplification of the present invention.  
XX  
SQ Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;  
Query Match 98.8%; Score 3323.4; DB 22; Length 3906;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;  
QY 1 GACTCCCGCATCCCAAAAGAACACCATGATGACGCAAAAAGAGATGGGCCCCCGAGC 60  
Db 563 GACTCCCGCATCCCAAAAGAACACCATGATGACGCAAAAAGAGATGGGCCCCCGAGC 622  
QY 61 CTCGTGCTGTGCTGCTGCTGCGCACTGTGTTCTCCCTGCTGGTGGAGCTCGGCCTTC 120  
Db 623 CTCGTGCTGTGCTGCTGCTGCGCACTGTGTTCTCCCTGCTGGTGGAGCTCGGCCTTC 682  
QY 121 CTGTGCGACCAACCGCTTGAAAGGCGAGGTTTCAGAGGACCGCAGGAAATCGGCCCCCAAC 180  
Db 683 CTGTGCGACCAACCGCTTGAAAGGCGAGGTTTCAGAGGACCGCAGGAAATCGGCCCCCAAC 742  
QY 181 ATCATCTCTGGTGTCTGACGAGCAGCAGGATGTGGAGCTGGTTCATCAGTGTATGAAC 240  
Db 743 ATCATCTCTGGTGTCTGACGAGCAGCAGGATGTGGAGCTGGTTCATCAGTGTATGAAC 802  
QY 241 AAGACCCCGCGCATCATGAGCAGGCGGGCGCGCATCTTCATCAACGCTTCGTGACCA 300  
Db 803 AAGACCCCGCGCATCATGAGCAGGCGGGCGCGCATCTTCATCAACGCTTCGTGACCA 862  
QY 301 CCCATGTGCTGCCCTTCAGGCTCTCCATCTCTCATGCGAAGTACGTCCCAACACCAAC 360  
Db 863 CCCATGTGCTGCCCTTCAGGCTCTCCATCTCTCATGCGAAGTACGTCCCAACACCAAC 922  
QY 361 ACCTACACCAACATGAGAACTGCTTCCTCGCCCTCTCGGACGCAACAGAGAGCGCG 420  
Db 923 ACCTACACCAACATGAGAACTGCTTCCTCGCCCTCTCGGACGCAACAGAGAGCGCG 982  
QY 421 ACCTTTGGGTGTACTCTCAATAGACTGGCTACCGACAGCTTTCTTCGGAGTATCTT 480  
Db 983 ACCTTTGGGTGTACTCTCAATAGACTGGCTACCGACAGCTTTCTTCGGAGTATCTT 1042  
QY 481 AATGAATACAAAGGCTCTCTACGTGCCACCCCGCTGGAAGGAGTGGGTCCGACTCTCTAAA 540  
Db 1043 AATGAATACAAAGGCTCTCTACGTGCCACCCCGCTGGAAGGAGTGGGTCCGACTCTCTAAA 1102

QY 541 AACTCCCGCTTTTATAACTACAGCTGTGTTCGAAACGGGCTGAAAGAAAGCAACGGCTCC 600  
DB 1103 AACTCCCGCTTTTATAACTACAGCTGTGTTCGAAACGGGCTGAAAGAAAGCAACGGCTCC 1162  
QY 601 GACTACTCCAGGATTAACCTACAGACCTTCATCAACCAATGACAGGTGAGCTTCTCCGC 660  
DB 1163 GACTACTCCAGGATTAACCTACAGACCTTCATCAACCAATGACAGGTGAGCTTCTTCGC 1222  
QY 661 ACCTCCAAAGAGATGATACCGGACAGGCGAGCTCCTCATGCTATCAGGCCATCAGCCGCC 720  
DB 1223 ACCTCCAAAGAGATGATACCGGACAGGCGAGCTCCTCATGCTATCAGGCCATCAGCCGCC 1282  
QY 721 CACGGCCCTGAGGATTCAGCCGCCACCAATATTCAGCCCTCTTCCCAAAACGCACTTCAGCAC 780  
DB 1283 CACGGCCCTGAGGATTCAGCCGCCACCAATATTCAGCCCTCTTCCCAAAACGCACTTCAGCAC 1342  
QY 781 ATCACCGCGAGCTACAACTACCGCGGCCCAACCGGACAAACACATGGATCATGCGCTACACG 840  
DB 1343 ATCACCGCGAGCTACAACTACCGCGGCCCAACCGGACAAACACATGGATCATGCGCTACACG 1402  
QY 841 GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCGAG 900  
DB 1403 GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCGAG 1462  
QY 901 ACCCTCATGTGCGTGAAGACTCCATGGAAGCTGAGACGATTTTACAAATGCTGTTGAGACGGC 960  
DB 1463 ACCCTCATGTGCGTGAAGACTCCATGGAAGCTGAGACGATTTTACAAATGCTGTTGAGACGGC 1522  
QY 961 GAGCTGACAAACAGCTACATCGTATACCGCGGACCAACGGTTTACCAATCGGCGAGTTT 1020  
DB 1523 GAGCTGACAAACAGCTACATCGTATACCGCGGACCAACGGTTTACCAATCGGCGAGTTT 1582  
QY 1021 GGCTGTGTAAGGGAAATTCATGCGCATATGAGTTTGAATCAGGGTCCCGTTTACAGTG 1080  
DB 1583 GGCTGTGTAAGGGAAATTCATGCGCATATGAGTTTGAATCAGGGTCCCGTTTACAGTG 1642  
QY 1081 AGGGCCCCAACTGGGAAGCGGCTGTCTGAATCCCCACATGCTCTCAACATTCAGCCTG 1140  
DB 1643 AGGGCCCCAACTGGGAAGCGGCTGTCTGAATCCCCACATGCTCTCAACATTCAGCCTG 1702  
QY 1141 GCCCCCAACCATCTCGACATTTGAGGACCTTGACATACCTTCGGGATGATGACGGGAAATCC 1200  
DB 1703 GCCCCCAACCATCTCGACATTTGAGGACCTTGACATACCTTCGGGATGATGACGGGAAATCC 1762  
QY 1201 ATCTCAAGCTGTGACAGGAGCGGCGGTGAATCGGTTTCACTTGAAGAGAGATG 1260  
DB 1763 ATCTCAAGCTGTGACAGGAGCGGCGGTGAATCGGTTTCACTTGAAGAGAGATG 1822  
QY 1261 AGGCTCTGGGGGACTCCTTTCTTGGTGGAGAGGCAAGCTGTACACAAGAGAGACAAAT 1320  
DB 1823 AGGCTCTGGGGGACTCCTTTCTTGGTGGAGAGGCAAGCTGTACACAAGAGAGACAAAT 1882  
QY 1321 GACAGGTGAGCGCCGAGGAGAGAACTTTCTGCCAAGTACAGCGGTGTGAAGACCTG 1380  
DB 1883 GACAAGGTGAGCGCCGAGGAGAGAACTTTCTGCCAAGTACAGCGGTGTGAAGACCTG 1942  
QY 1381 TGTCAAGCTGTGAGTACACAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTG 1440  
DB 1943 TGTCAAGCTGTGAGTACACAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTG 2002  
QY 1441 GAGGACGCCACGGGAAAGCTGAAGCTGCATTAAGTGAAGGGGCCCATCGGCTGGCGGC 1500  
DB 2003 GAGGACGCCACGGGAAAGCTGAAGCTGCATTAAGTGAAGGGGCCCATCGGCTGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCCAACTCTGCCCCAAGTACTACGGGCGAGGCGAGCGGCTGCACC 1560  
DB 2063 AGCAGAGCCCTCTCCAACTCTGCCCCAAGTACTACGGGCGAGGCGAGCGGCTGCACC 2122  
QY 1561 TGTGACAGCGGGGACTACAAGCTCAGCTCGCGGACCGCGGAAAAAATCTTTCAAGAAAG 1620  
DB 2123 TGTGACAGCGGGGACTACAAGCTCAGCTCGCGGACCGCGGAAAAAATCTTTCAAGAAAG 2182

QY 1621 AAGTACAAGGCCAGCTATGTTCGCACTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680  
DB 2183 AAGTACAAGGCCAGCTATGTTCGCACTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2242  
QY 1681 GGCAGGGTGTACACCTAGGCTGGGTGATGCGGCCCGCCGCAACCTCCACCAAGCGG 1740  
DB 2243 GGCAGGGTGTACACCTAGGCTGGGTGATGCGGCCCGCCGCAACCTCCACCAAGCGG 2302  
QY 1741 CACTGGCCAGGGGCCCTCTGAGGACCAAGATGACAAGATGTTGGGGACTTCAGTGGCACT 1800  
DB 2303 CACTGGCCAGGGGCCCTCTGAGGACCAAGATGACAAGATGTTGGGGACTTCAGTGGCACT 2362  
QY 1801 GAGAGGCTTCCCGACTACTCAGCGGCCAACCCTATTAAGTGAACATCGGTGCTACATC 1860  
DB 2363 GAGAGGCTTCCCGACTACTCAGCGGCCAACCCTATTAAGTGAACATCGGTGCTACATC 2422  
QY 1861 CTAGAGAAACACACAGTCAGTGTGACCTGAGACCTGTACAAGTCCCTGCAAGGCTCGAAA 1920  
DB 2423 CTAGAGAAACACACAGTCAGTGTGACCTGAGACCTGTACAAGTCCCTGCAAGGCTCGAAA 2482  
QY 1921 GACCAAAAGCTGCAATCGACCAAGATTTGAAACCTCTGCAAGAACAAATTAAGAACCTG 1980  
DB 2483 GACCAAAAGCTGCAATCGACCAAGATTTGAAACCTCTGCAAGAACAAATTAAGAACCTG 2542  
QY 1981 AGGGAAGCTCCGAGTCACTGAAAGAAAGGCGGCCAGAAAGATGTGACTGTCAAAAAATC 2040  
DB 2543 AGGGAAGCTCCGAGTCACTGAAAGAAAGGCGGCCAGAAAGATGTGACTGTCAAAAAATC 2602  
QY 2041 AGCTTACACACCCAGGACCAAGGCCCTCAAGCACAGAGGCTCCAGTGTGCAATCTTTT 2100  
DB 2603 AGCTTACACACCCAGGACCAAGGCCCTCAAGCACAGAGGCTCCAGTGTGCAATCTTTT 2662  
QY 2101 AGGAGGGCTCTCAAGAGAGGACAGGTGTGCTCTGCGGAGCAGAGCGCAAGAG 2160  
DB 2663 AGGAGGGCTCTCAAGAGAGGACAGGTGTGCTCTGCGGAGCAGAGCGCAAGAG 2722  
QY 2161 AAATCTCGCAAGCTGTCTCAAGCGCTCGAGAACCAACGACACAGCTGCGAGGCTC 2220  
DB 2723 AAATCTCGCAAGCTGTCTCAAGCGCTCGAGAACCAACGACACAGCTGCGAGGCTC 2782  
QY 2221 ACGTGTCTTCAACCAAGCAACCAAGCTGCGAGACCGCGCTTTCTGGACACTGGGGCT 2280  
DB 2783 ACSTGTCTTCAACCAAGCAACCAAGCTGCGAGACCGCGCTTTCTGGACACTGGGGCT 2842  
QY 2281 TTCTGTGCTTGCACACGCGCAACCAATTAACAGCTACTGTGTCATGAGGACCAATGAG 2340  
DB 2843 TTCTGTGCTTGCACACGCGCAACCAATTAACAGCTACTGTGTCATGAGGACCAATGAG 2902  
QY 2341 ACTCAAAATTTCTTCTGTGAAATTTGCAATCTGGCTTCTTAGAGTACTTTGATCTCAAC 2400  
DB 2903 ACTCAAAATTTCTTCTGTGAAATTTGCAATCTGGCTTCTTAGAGTACTTTGATCTCAAC 2962  
QY 2401 ACAGACCCCTTACAGCTGATGATGCAAGTGAACACACTGAGCAGGGATGCTTCAACAG 2460  
DB 2963 ACAGACCCCTTACAGCTGATGATGCAAGTGAACACACTGAGCAGGGATGCTTCAACAG 3022  
QY 2461 CTACACGTTACAGCTCATGAGGCTGAGGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGG 2520  
DB 3023 CTACACGTTACAGCTCATGAGGCTGAGGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGG 3082  
QY 2521 ACTGGAACATGAGCTTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGCAGTTT 2580  
DB 3083 ACTGGAACATGAGCTTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGCAGTTT 3133  
QY 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCAGCTGGGACCAATGTGG 2640  
DB 3134 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCAGCTGGGACCAATGTGG 3193  
QY 2641 GAAGCTGGGAGGTTTGAAGAAACACAGAGGTGAGCTTCCAAACACATAGAGGCAATCACC 2700  
DB 3194 GAAGCTGGGAGGTTTGAAGAAACACAGAGGTGAGCTTCCAAACACATAGAGGCAATCACC 3253  
QY 2701 TGACTGCAAGGCAATGAAGAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760

Db 3254 TGACTGCACAGGCAATGAAAAACCATGTGGTGTATTCAGCAGACCTGTCTATTGGCC 3313  
QY 2761 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATCTTGAGGATTAAC 2820  
Db 3314 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATCTTGAGGATTAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAACTTCAAGGAAGTCCATTTTGGCCCTGCTTTTGGTATTA 2880  
Db 3374 CAGCAGGAGCAGAGATAACTTCAAGGAAGTCCATTTTGGCCCTGCTTTTGGTATTA 3433  
QY 2881 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAAGTCAACCTTCC 2940  
Db 3434 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAAGTCAACCTTCC 3493  
QY 2941 CCAGAAGCTCACAAGGAAAAACGAGAGAGCGAGAGATTTCCCTTGGAAATTTCT 3000  
Db 3494 CCAGAAGCTCACAAGGAAAAACGAGAGAGCGAGAGATTTCCCTTGGAAATTTCT 3553  
QY 3001 CCCAAGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGAGTCTCTGTTCTAAA 3060  
Db 3554 CCCAAGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGAGTCTCTGTTCTAAA 3613  
QY 3061 TCCTCTATTCTTTTGGTTTGTCAAAAGAGAACTTAAGAGAGGAGAGGCAAGCT 3120  
Db 3614 TCCTCTATTCTTTTGGTTTGTCAAAAGAGAACTTAAGAGAGGAGAGGCAAGCT 3673  
QY 3121 GGAGAGCTGAAACAGTGCAGAGAGCGTTTGAATAGTCAAGTCAAGCAAAAGAGATGA 3180  
Db 3674 GGAGAGCTGAAACAGTGCAGAGAGCGTTTGAATAGTCAAGTCAAGCAAAAGAGATGA 3733  
QY 3181 CATTTACCTAGCACTATAACCTGCTTCCCTCTGAAGAACTGCTTCATTGTATATAT 3240  
Db 3734 CATTTACCTAGCACTATAACCTGCTTCCCTCTGAAGAACTGCTTCATTGTATATAT 3793  
QY 3241 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAAGAAGAAATCCCAA 3300  
Db 3794 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAAGAAGAAATCCCAA 3853  
QY 3301 TTTTGAAGAGTGTGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3353  
Db 3854 TTTTGAAGAGTGTGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3906

RESULT 11  
ABL95597  
ID ABL95597 standard; cDNA; 3906 BP.  
XX  
AC ABL95597;  
XX  
XX  
DT 19-JUL-2002 (first entry)  
XX  
DE Human angiogenesis related cDNA PRO1120 SEQ ID NO: 73.  
XX  
KW Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;  
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;  
KW cardiant; cytosstatic; antiangiogenic; hypotensive; vulnerary;  
KW antiarteriosclerotic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200208284-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 09-JUL-2001; 2001WO-US21735.  
XX  
PR 20-JUL-2000; 2000US-219556P.  
PR 25-JUL-2000; 2000US-220624P.  
PR 25-JUL-2000; 2000US-220664P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 02-AUG-2000; 2000US-222695P.  
PR 17-AUG-2000; 2000US-0643657.

PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 07-SEP-2000; 2000US-230978P.  
PR 15-SEP-2000; 2000US-000000P.  
PR 18-SEP-2000; 2000US-064610.  
PR 18-SEP-2000; 2000US-066530.  
PR 24-OCT-2000; 2000US-242922P.  
PR 08-NOV-2000; 2000US-0709238.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000US-0747259.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 22-JAN-2001; 2001US-0767609.  
PR 28-FEB-2001; 2001US-0796498.  
PR 01-MAR-2001; 2001WO-US06520.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808889.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 10-MAY-2001; 2001US-0854280.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 30-MAY-2001; 2001US-0870574.  
PR 30-MAY-2001; 2001WO-US17443.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 28-JUN-2001; 2001WO-US00000.  
XX (GETH ) GENENTECH INC.  
PA (BAKE/) BAKER K P.  
PA (FERR/) FERRARA N.  
PA (GERB/) GERBER H.  
PA (GERR/) GERRITSEN M E.  
PA (GODD/) GODDARD A. J.  
PA (GODO/) GODOWSKI P. J.  
PA (GURN/) GURNEY A L.  
PA (HILL/) HILLAN K J.  
PA (MARS/) MARSTERS S A.  
PA (PANJ/) PAN J.  
PA (PAON/) PAONI N F.  
PA (STEP/) STEPHAN J F.  
PA (WATA/) WATANABE C K.  
PA (WILL/) WILLIAMS P M.  
PA (WOOD/) WOOD W I.  
XX  
PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A,  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF,  
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
XX  
DR WPI; 2002-171999/22.  
DR P-PSDB; ABB95459.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal -  
XX  
XX Claim 1; Fig 73; 567pp; English.  
XX  
XX The present invention provides the protein and coding sequences of human  
XX PRO proteins. These are useful for treating or diagnosing a  
XX cardiovascular, endothelial or angiogenic disorder, including cardiac  
XX hypertrophy, trauma, cancer, age-related macular degeneration,  
XX atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
XX angina, myocardial infarctions, thrombopilebitis, lymphangitis, tumour  
XX angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
XX healing. The present sequence is a coding sequence of the invention.  
XX  
XX Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;









Db	2183	AAGTAAAGGCCAGCTATGTCGCCAGATCGCTCCATCCGCTAGTGGCCATCGAGGTGGAC	2244
Qy	1681	GGCAGGCTGTACCACTGAGCCCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG	1740
Db	2243	GGCAGGCTGTACCACTGAGCCCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG	2302
Qy	1741	CAC TGCCAGGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACNTTCAGTGGCACT	1800
Db	2303	CAC TGCCAGGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACNTTCAGTGGCACT	2362
Qy	1801	GGAGGCCTTCCGACTACTCAGCGCCCAACCCCATTAAGTGTGACATCCTGGTGTACATC	1860
Db	2363	GGAGGCCTTCCGACTACTCAGCGCCCAACCCCATTAAGTGTGACATCCTGGTGTACATC	2422
Qy	1861	CTAGAGAACACACACAGTGCAGTGTGACCTGGACCTGTACAAGTCCCTGAGGCTGGAAA	1920
Db	2423	CTAGAGAACACACACAGTGCAGTGTGACCTGGACCTGTACAAGTCCCTGAGGCTGGAAA	2482
Qy	1921	GACCACAAGCTGCACATCGACCAGATTTGAAACCTCGACAGAACAAATTAAGAACCTTG	1980
Db	2483	GACCACAAGCTGCACATCGACCAGATTTGAAACCTCGACAGAACAAATTAAGAACCTTG	2542
Qy	1981	AGGGAAGTCCGAGGTCACTGTGAAGAAAAGCGGCGCAGAGAAATGTGACTGTCAAAAATC	2040
Db	2543	AGGGAAGTCCGAGGTCACTGTGAAGAAAAGCGGCGCAGAGAAATGTGACTGTCAAAAATC	2602
Qy	2041	AGCTTACCACACCAGACAAAGCGCCTCAAGCACACAGAGGCTCCAGTGTGCATCTCTTTC	2100
Db	2603	AGCTTACCACACCAGACAAAGCGCCTCAAGCACACAGAGGCTCCAGTGTGCATCTCTTTC	2662
Qy	2101	AGGAAGGCCTGCAAGAGAAAGGACAAGGTGTGCTGTTCGGGAGCAGAAAGCCCAAGAAG	2160
Db	2663	AGGAAGGCCTGCAAGAGAAAGGACAAGGTGTGCTGTTCGGGAGCAGAAAGCCCAAGAAG	2722
Qy	2161	AACTCCGCAAGCTGCTCAAGCGCCTGCAGAACACGACACGTCGACATGCCAGGCTTC	2220
Db	2723	AACTCCGCAAGCTGCTCAAGCGCCTGCAGAACACGACACGTCGACATGCCAGGCTTC	2782
Qy	2221	ACGTGCTTCCACCAACACGACCTGGCAGACGGCGCCTTTCTGGACACATCGGGGCGCT	2280
Db	2783	ACGTGCTTCCACCAACACGACCTGGCAGACGGCGCCTTTCTGGACACATCGGGGCGCT	2842
Qy	2281	TTCTGTGCTGTCACAGCGGCGCAATPAAACAGTACGTGTGCTGTGAGGACCATCAATGAG	2340
Db	2843	TTCTGTGCTGTCACAGCGGCGCAATPAAACAGTACGTGTGCTGTGAGGACCATCAATGAG	2902
Qy	2341	ACTCACAAATTTCTCTCTGTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCAAC	2400
Db	2903	ACTCACAAATTTCTCTCTGTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCAAC	2962
Qy	2401	ACAGACCCCTACCAAGCTGTGAATGCAAGTGAAACACACTGACAGGGATGTCTCAACCAAG	2460
Db	2963	ACAGACCCCTACCAAGCTGTGAATGCAAGTGAAACACACTGACAGGGATGTCTCAACCAAG	3022
Qy	2461	CTACACCTPACAGCTCATGGAGCTGAGGAGCTGCAGGGTTTACAGCAGTGTAAACCCCGG	2520
Db	3023	CTACACCTPACAGCTCATGGAGCTGAGGAGCTGCAGGGTTTACAGCAGTGTAAACCCCGG	3082
Qy	2521	ACTCGAAACATGGACCTGGGACCTTTAAAGATGGAGAGCTATGAGCAATACAGCGCAGTTT	2580
Db	3083	ACTCGAAACATGGACCTG-----GATGGAGGAAGCTATGAGCAATACAGCGCAGTTT	3133
Qy	2581	CAGGCTCGAAAGTGGCCAGAAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG	2640
Db	3134	CAGGCTCGAAAGTGGCCAGAAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG	3193
Qy	2641	GAAGGCTTGGGAAGGTTAAGAAAACACAGAGGTGGAACCTCCAAAAACATAGAGGCATCACC	2700
Db	3194	GAAGGCTTGGGAAGGTTAAGAAAACACAGAGGTGGAACCTCCAAAAACATAGAGGCATCACC	3253
Qy	2701	TGACTGCACAGGCAATGAAAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC	2760
Db	3254	TGACTGCACAGGCAATGAAAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC	3313

QY	2761	AGGAGGCTGAGAAAGCAGCAGCCACTCTCAGTCAACATGACAGATTCGGAGGATAAC	2822
DB	3314	AGGAGGCTGAGAAAGCAGCAGCCACTCTCAGTCAACATGACAGATTCGGAGGATAAC	3373
QY	2821	CAGCAGGAGCAGAGATAACTTCAGGAAAGTCCATTTTTCGCCCTGCTTTTTCGTTTGGATTA	2880
DB	3374	CAGCAGGAGCAGAGATAACTTCAGGAAAGTCCATTTTTCGCCCTGCTTTTTCGTTTGGATTA	3433
QY	2881	TACCTCACCGCTGCGACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC	2940
DB	3434	TACCTCACCGCTGCGACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC	3493
QY	2941	CCGAAAGCTCACAAAGGAAAAACGAGAGAGCGAGCGAGAGAGATTTTCCTTGGAAATTTCT	3000
DB	3494	CCGAAAGCTCACAAAGGAAAAACGAGAGAGCGAGCGAGAGAGATTTTCCTTGGAAATTTCT	3553
QY	3001	CCCAAGGCGGAAGTCAATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTCTAAA	3066
DB	3554	CCCAAGGCGGAAGTCAATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTCTAAA	3613
QY	3061	TCCTCTTATTCCTTTTGGTTTGTCAAAAGGAACTAAGAAAGCAGGACAGAGCAACGT	3120
DB	3614	TCCTCTTATTCCTTTTGGTTTGTCAAAAGGAACTAAGAAAGCAGGACAGAGCAACGT	3673
QY	3121	GGAGAGCTGAAAACAGTCGAGAGACGTTTTCACATGAGTCAGTAGCACAAAAGAGATGA	3180
DB	3674	GGAGAGCTGAAAACAGTCGAGAGACGTTTTCACATGAGTCAGTAGCACAAAAGAGATGA	3733
QY	3181	CATTTCCTAGCACTATAAAACCCCTGGTTGCTCTGAAAGAACTGCGCTCATTTGTATATAT	3240
DB	3734	CATTTCCTAGCACTATAAAACCCCTGGTTGCTCTGAAAGAACTGCGCTCATTTGTATATAT	3793
QY	3241	GTGACTATTTACATGTAATCAACATAGGAACTTTTAGGGGAACTTAATAAGAAATCCCAA	3300
DB	3794	GTGACTATTTACATGTAATCAACATAGGAACTTTTAGGGGAACTTAATAAGAAATCCCAA	3853
QY	3301	TTTTTCAGAGTGGTGGTGTCATTAACCGCTCTGTGGCCAGTGTAAAAAGAAAAA	3353
DB	3854	TTTTTCAGAGTGGTGGTGTCATTAACCGCTCTGTGGCCAGTGTAAAAAGAAAAA	3906
RESULT 13			
ACA03612			
ID	ACA03612 standard; cDNA; 3906 BP.		
AC	ACA03612;		
DT	23-MAY-2003 (first entry)		
DE	cDNA encoding human PRO polypeptide #10.		
KW	Human; PRO polypeptide; secreted and transmembrane protein;		
KW	tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;		
KW	differentiation; chondrocyte; tumour; genetic disorder;		
KW	cytostatic; gene; ss.		
OS	Homo sapiens.		
PN	US2003036180-A1.		
XX	20-FEB-2003.		
PF	09-MAY-2002; 2002US-0143114.		
XX	31-MAR-1997; 97WO-US05230.		
PR	12-JUN-1998; 98WO-US12456.		
PR	14-JUN-1998; 98WO-US14552.		
PR	28-AUG-1998; 98WO-US17888.		
PR	10-SEP-1998; 98WO-US18844.		
PR	14-SEP-1998; 98WO-US19093.		
PR	14-SEP-1998; 98WO-US19094.		
PR	14-SEP-1998; 98WO-US19177.		







PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14045.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US2031.  
PR 23-AUG-2000; 2000WO-US2352.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 18-MAY-2001; 2001US-0854280.  
PR 25-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.

(GETH ) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI: 2003-331925/31.  
P-PSDB; ABU66855.

New secreted and transmembrane nucleic acids and polypeptides,  
designated as PRO, useful for treating inflammation, organ failure,  
atherosclerosis, cardiac injury, infertility, birth defects, premature  
aging, AIDS, or cancer

Claim 2; Fig 19; 659pp; English.

The invention relates to an isolated nucleic acid comprising, or which is  
at least 80% identical to, or the full-length coding sequence of, any of  
the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
(one of 275 secreted or transmembrane proteins). The nucleic acid  
further comprises the full-length coding sequence of the DNA deposited  
under American Type Culture Collection (ATCC) accession number in a list  
given in the specification. Also included are vectors and host  
cells for producing PRO proteins, PRO fusion proteins, anti-PRO  
antibodies, PRO extracellular domains and mature sequences, methods  
of detecting PRO proteins, methods for stimulating the release of  
TNF-alpha (tumour necrosis factor alpha) from human blood,  
(and the proliferation of differentiation of chondrocyte cells, the

CC proliferation of, or gene expression in pericyte cells, the release or  
CC proteoglycans from cartilage, proliferation of inner ear utricular  
CC supporting cells, the proliferation of T-lymphocyte cells, the release  
CC of a cytokine from peripheral blood mononuclear cells (PBMC), or the  
CC proliferation of endothelial cells), a method for modulating the uptake  
CC of glucose or free fatty acid (FFA) by skeletal muscle cells,  
CC a method for inhibiting the binding of A-peptide to factor VIIA,  
CC or the differentiation of adipocyte cells, a method for detecting the  
CC presence of a tumour in a mammal and an oligonucleotide probe derived  
CC from any of the nucleotide sequences cited above. The nucleic acids and  
CC polypeptides are useful for treating inflammatory diseases, organ  
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,  
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or  
CC diabetic complications. The nucleic acids are useful as hybridisation  
CC probes, in chromosome and gene mapping, and in generating antisense RNA  
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,  
CC biosensors or bioreactors. Both are useful in tissue typing.  
CC The present sequence encodes a PRO protein of the invention.

XX Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;

Query Match 98.8%; Score 3323.4; DB 25; Length 3906;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 GACTCCGCATCCCAAAGAGAGACACAGATCAGCAAAAAGAGAGATGGCCCCCGAGC 60  
DB 563 GACTCCGCATCCCAAAGAGAGACACAGATCAGCAAAAAGAGAGATGGCCCCCGAGC 622  
QY 61 CTCGTGCTGTGCTGCTGCCCAACTGTCTTCTCCCTGCTGGGTGGAAGCTCGGCCCTTC 120  
DB 623 CTCGTGCTGTGCTGCTGCCCAACTGTCTTCTCCCTGCTGGGTGGAAGCTCGGCCCTTC 682  
QY 121 CTGTGCAACACCGCTCTGAAG 180  
DB 683 CTGTGCAACACCGCTCTGAAG 742  
QY 181 ATCATCTGCTGCTGCTGCAAG 240  
DB 743 ATCATCTGCTGCTGCTGCAAG 802  
QY 241 AAGACCCCGCGCATCATGAG 300  
DB 803 AAGACCCCGCGCATCATGAG 862  
QY 301 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
DB 863 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922  
QY 361 ACTATACCAACAATGAG 420  
DB 923 ACTATACCAACAATGAG 982  
QY 421 ACTTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 983 ACTTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042  
QY 481 AATGAATACCAAG 540  
DB 1043 AATGAATACCAAG 1102  
QY 541 AACTCCCGCTTTTATAAATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 1103 AACTCCCGCTTTTATAAATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162  
QY 601 GACTACTCCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 1163 GACTACTCCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1222  
QY 661 AGCTTCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 1223 AGCTTCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1282

QY 721 CACGCGCTGAGGATTCAGCCCCCAATATTTACGCGCTCTTCCAAACGATCTCAGCAC 780  
DB 1283 CACGCGCTGAGGATTCAGCCCCCAATATTTACGCGCTCTTCCAAACGATCTCAGCAC 1342  
QY 781 ATCAGCGCGAGGTAACTACGCGGCCAACCGGACAAAACACTGGATCCTGCGGTACAGC 840  
DB 1343 ATCAGCGCGAGGTAACTACGCGGCCAACCGGACAAAACACTGGATCCTGCGGTACAGC 1402  
QY 841 GGGCCCATGAAAGCCCATCCACATGGAATTTACCAACATGCTCCAGCGGAAAGCGTTTCAG 900  
DB 1403 GGGCCCATGAAAGCCCATCCACATGGAATTTACCAACATGCTCCAGCGGAAAGCGTTTCAG 1462  
QY 901 ACCCTCATGTGCTGGAGCGATCCATGAGAGCGATTTTACAACTGCTGGTTGAGACGGC 960  
DB 1463 ACCCTCATGTGCTGGAGCGATCCATGAGAGCGATTTTACAACTGCTGGTTGAGACGGC 1522  
QY 961 GAGCTGGGACAAACAGCTACATCGTATACACCGCGAGACCAACGCTTACCACTCGCGCAGTTT 1020  
DB 1523 GAGCTGGGACAAACAGCTACATCGTATACACCGCGAGACCAACGCTTACCACTCGCGCAGTTT 1582  
QY 1021 GGCCTGGTAAAGGAAATCCATGCGCATATGAGTTTGAATCAGGGTCCCGTTCTACGTG 1080  
DB 1583 GGCCTGGTAAAGGAAATCCATGCGCATATGAGTTTGAATCAGGGTCCCGTTCTACGTG 1642  
QY 1081 AGGCGGCCCAACGCTGAAAGCGGCTGTCTGAATCCCACTCCTCCTCAACATTTGACTG 1140  
DB 1643 AGGCGGCCCAACGCTGAAAGCGGCTGTCTGAATCCCACTCCTCCTCAACATTTGACTG 1702  
QY 1141 GCCCCACCATCTCTGGACATTTGAGGCTGGACATACCTGCGGATATGACGGGAAATCC 1200  
DB 1703 GCCCCACCATCTCTGGACATTTGAGGCTGGACATACCTGCGGATATGACGGGAAATCC 1762  
QY 1201 ATCTCTAAGCTCTGGACACGAGCGGCGGTGATCGTTTCTCTTGAAGAAGAGATG 1260  
DB 1763 ATCTCTAAGCTCTGGACACGAGCGGCGGTGATCGTTTCTCTTGAAGAAGAGATG 1822  
QY 1261 AGGCGTCTGGCGGAACTCTCTTTCTTGGTGGAGAGGCAAGCTGTCTACACAAGAGAGACAAT 1320  
DB 1823 AGGCGTCTGGCGGAACTCTCTTTCTTGGTGGAGAGGCAAGCTGTCTACACAAGAGAGACAAT 1882  
QY 1321 GACAAGTGGAGCGCCGAGAGAGAACTTTCTGCCCAAGTACCAGCGTGTCAAGACCTG 1380  
DB 1883 GACAAGTGGAGCGCCGAGAGAGAACTTTCTGCCCAAGTACCAGCGTGTCAAGACCTG 1942  
QY 1381 TGTACAGCTGTGAGTACACAGCGCGGTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG 1440  
DB 1943 TGTACAGCTGTGAGTACACAGCGCGGTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG 2002  
QY 1441 GAGGCGCCACGCGGGAAGCTGAAGCTGCAATAAGTGCAGAGGCGCCCATCGCGGTGGCGGC 1500  
DB 2003 GAGGCGCCACGCGGGAAGCTGAAGCTGCAATAAGTGCAGAGGCGCCCATCGCGGTGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTAGGCGAGGCGAGGCGCTGCAACC 1560  
DB 2063 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTAGGCGAGGCGAGGCGCTGCAACC 2122  
QY 1561 TGTGACAGCGGGAATAAAGCTCAGCTGAGCTGGCGGACGCGGAAAAAACTCTTCAAGAAG 1620  
DB 2123 TGTGACAGCGGGAATAAAGCTCAGCTGAGCTGGCGGACGCGGAAAAAACTCTTCAAGAAG 2182  
QY 1621 AAGTACAAGCGCAGCTATCTCGCAGTGCCTCCATCCGCTCAGTCGCCATCGAGGTGGAC 1680  
DB 2183 AAGTACAAGCGCAGCTATCTCGCAGTGCCTCCATCCGCTCAGTCGCCATCGAGGTGGAC 2242  
QY 1681 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGGCCAGCCCCGAAAACTCACCAGCGG 1740  
DB 2243 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGGCCAGCCCCGAAAACTCACCAGCGG 2302  
QY 1741 CACTGCGCAGGGCCCCCTGAGGACCAAGATGAGAGGATGGTGGGACTTCTAGTGGCACT 1800  
DB 2303 CACTGCGCAGGGCCCCCTGAGGACCAAGATGAGAGGATGGTGGGACTTCTAGTGGCACT 2362  
QY 1801 GGAGGGCTTCCCGACTACTCAGCGCGCCAAACCCCAATTAAGGTGACACATCGGTGTACATC 1860

DB 2363 GGAGGGCTTCCCGACTACTCAGCGCGCCAAACCCCAATTAAGGTGACACATCGGTGTACATC 2422  
QY 1861 CTAGAGAAACGACACAGTTCAGTGTGACCTGACCTGTACAAAGTCCCTTCAGGCGCTGAAA 1920  
DB 2423 CTAGAGAAACGACACAGTTCAGTGTGACCTGACCTGTACAAAGTCCCTTCAGGCGCTGAAA 2482  
QY 1921 GACCAACAAGTGTGCAATCGACCAACAGATTGAACCTCTGACGAAACAAAATTAAGAACCTG 1980  
DB 2483 GACCAACAAGTGTGCAATCGACCAACAGATTGAACCTCTGACGAAACAAAATTAAGAACCTG 2542  
QY 1981 AGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGCCGAGAAAGATGTGACTGTCACAAAATC 2040  
DB 2543 AGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGCCGAGAAAGATGTGACTGTCACAAAATC 2602  
QY 2041 AGCTACCAACCCAGCAACAAAGCGCCTCAAGCAACAGAGGCTCCAGTCTGTGATCCTTTT 2100  
DB 2603 AGCTACCAACCCAGCAACAAAGCGCCTCAAGCAACAGAGGCTCCAGTCTGTGATCCTTTT 2662  
QY 2101 AGGAAGGCGCTCGAAGAGAGAGCAAGGTGTGGCTGTTCGCGGAGCAGAAAGCGCAAGAG 2160  
DB 2663 AGGAAGGCGCTCGAAGAGAGAGCAAGGTGTGGCTGTTCGCGGAGCAGAAAGCGCAAGAG 2722  
QY 2161 AAACTCCGCAAGCTGCTCAAGCGCCTGACAGAAACAAACGACAGCTGACGATGCCAGCCTC 2220  
DB 2723 AAACTCCGCAAGCTGCTCAAGCGCCTGACAGAAACAAACGACAGCTGACGATGCCAGCCTC 2782  
QY 2221 AGTGTCTTCAACCAACGACCAAGCTGGCAGACGCGCCTTTCTGGACACTGGGGCT 2280  
DB 2783 AGTGTCTTCAACCAACGACCAAGCTGGCAGACGCGCCTTTCTGGACACTGGGGCT 2842  
QY 2281 TTCTGTGCTGCAACCAAGCGCCCAACAAATAACAGTACTGTGTGATGAGGACCAATCAATGAG 2340  
DB 2843 TTCTGTGCTGCAACCAAGCGCCCAACAAATAACAGTACTGTGTGATGAGGACCAATCAATGAG 2902  
QY 2341 ACTCAAAATTTCTCTTCTGTGAAATTTGCAACTGCTTCCTAGAGTACTTTGATCTCAAC 2400  
DB 2903 ACTCAAAATTTCTCTTCTGTGAAATTTGCAACTGCTTCCTAGAGTACTTTGATCTCAAC 2962  
QY 2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACACTGGAACAGGATGTCTCAACCAG 2460  
DB 2963 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACACTGGAACAGGATGTCTCAACCAG 3022  
QY 2461 CTACACAGTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG 2520  
DB 3023 CTACACAGTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG 3082  
QY 2521 ACTCGAAACATGAGCTGGGACTTTAAAGATGAGGAAGCTATGAGCAATAACAGGCAATTT 2580  
DB 3083 ACTCGAAACATGAGCTGGGACTTTAAAGATGAGGAAGCTATGAGCAATAACAGGCAATTT 3133  
QY 2581 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAAATCACTGGGAACAATGTGG 2640  
DB 3134 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAAATCACTGGGAACAATGTGG 3193  
QY 2641 GAAGGCTGGGAAGTTTAAAGAAACACAGAGGTGGACCTCCAAACACATAGAGCATCAC 2700  
DB 3194 GAAGGCTGGGAAGTTTAAAGAAACACAGAGGTGGACCTCCAAACACATAGAGCATCAC 3253  
QY 2701 TGAATGACAGGCAATGAAGAAACCAATGTGGGTGATTTTCAGCAGACCTGTGTATTGGCC 2760  
DB 3254 TGAATGACAGGCAATGAAGAAACCAATGTGGGTGATTTTCAGCAGACCTGTGTATTGGCC 3313  
QY 2761 AGGAGGCTGAGAAACACAGCAGCAGCTCTCAAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
DB 3314 AGGAGGCTGAGAAACACAGCAGCAGCTCTCAAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAAATCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGATTA 2880  
DB 3374 CAGCAGGAGCAGAGATAAATCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGATTA 3433  
QY 2881 TACTCAGCAGCTGCAAAAATGCAATTTTGTATCAAAAAGTCAACCAATCAACCTCC 2940



Db 3434 TACCTCACCGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCACCACTAACCTCC 3493  
Qy 2941 CCAGAAGCTCAAAAGGAAAACGGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3000  
Db 3494 CCAGAGCTCAAAAGGAAAACGGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3553  
Qy 3001 CCCAAGGCGAAAGTCATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3060  
Db 3554 CCCAAGGCGAAAGTCATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3613  
Qy 3061 TCCTCTTATTTCTTTTGTTCCTCAAGAGAGAACTTAAGACAGACAGACGACGACGTC 3120  
Db 3614 TCCTCTTATTTCTTTTGTTCCTCAAGAGAGAACTTAAGACAGACAGACGACGACGTC 3673  
Qy 3121 GGAGAGCTGAAAACAGTCGACAGACGTTTGACAATGAGTCAGTAGCACAAGAGAGATGA 3180  
Db 3674 GGAGAGCTGAAAACAGTCGACAGACGTTTGACAATGAGTCAGTAGCACAAGAGAGATGA 3733  
Qy 3181 CATTACTAGCATATATAACCTGGTTCCTCTGAAGAACTGCTTCATTGTATATAT 3240  
Db 3734 CATTACTAGCATATATAACCTGGTTCCTCTGAAGAACTGCTTCATTGTATATAT 3793  
Qy 3241 GTGACTATTTACATGTAATCAACATGGAACTTTAGGGGAACCTTAATAGAAATCCCA 3300  
Db 3794 GTGACTATTTACATGTAATCAACATGGAACTTTAGGGGAACCTTAATAGAAATCCCA 3853  
Qy 3301 TTTTCAGAGTGGTGTGTCTAATAAAGCTCTGTGGCCAGTGTAAAGAAAAA 3353  
Db 3854 TTTTCAGAGTGGTGTGTCTAATAAAGCTCTGTGGCCAGTGTAAAGAAAAA 3906

## RESULT 15

ABX89150  
ID ABX89150 standard; cDNA; 3906 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

DNA encoding novel secreted and transmembrane protein PRO1120.

Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
cardiac insufficiency disorder; cancer; tumour; immune response;  
adrenal cortical capillary endothelial growth; c-fos induction;  
vascular endothelial growth factor inhibition; VEGF inhibition;  
endothelial cell growth inhibitor; T-lymphocytes stimulation;  
retinal neurons cell survival; rod photoreceptor cell survival;  
retinal disorder; retinitis pigmentosa; kidney disorder;  
mamalian kidney mesangial cell proliferation; Berger disease;  
dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
chondrocyte redifferentiation; sports injury; arthritis; gene; ss.

Homo sapiens.

OS  
XX  
FN US2003017563-A1.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 10-MAR-1999; 99WO-US05190.  
PR 20-APR-1999; 99WO-US08615.  
PR 14-MAY-1999; 99WO-US10733.  
PR 02-JUN-1999; 99WO-US12252.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 30-NOV-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 22-DEC-1999; 99WO-US30720.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796499.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 10-MAY-2001; 2001US-0854280.



QY 1201 ATCCTAAGCTGCTGGACACGAGCGCGCGTGAATCGGTTTTCACTTGAAAAAGAGATG 1260  
Db 1763 ATCCTCAAGCTGCTGGACACGAGCGCGCGTGAATCGGTTTTCACTTGAAAAAGAGATG 1822  
QY 1261 AGGCTCTGGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACACAGAGAGCAAT 1320  
Db 1823 AGGCTCTGGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACACAGAGAGCAAT 1882  
QY 1321 GACAAGGTGGACGCGCCAGGAGGAGAACTTTCTGCGCCAAAGTACCAGCGTGTGAAGGACCTG 1380  
Db 1883 GACAAGGTGGACGCGCCAGGAGGAGAACTTTCTGCGCCAAAGTACCAGCGTGTGAAGGACCTG 1942  
QY 1381 TGTGAGCGTGTGATGATACAGACGCGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440  
Db 1943 TGTGAGCGTGTGATGATACAGACGCGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2002  
QY 1441 GAGGAGCGCACGGGAGAGCTGAAGCTGATGAAGTGAAGGCGGCCCATGCGGCTGGGCGGC 1500  
Db 2003 GAGGAGCGCACGGGAGAGCTGAAGCTGATGAAGTGAAGGCGGCCCATGCGGCTGGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCCAACTCGTGCCTCCCAAGTACTA CGGCGAGGCGCAGAGGCTTGCAAC 1560  
Db 2063 AGCAGAGCCCTCTCCAACTCGTGCCTCCCAAGTACTA CGGCGAGGCGCAGAGGCTTGCAAC 2122  
QY 1561 TGTGACAGCGGGGACTACAAGCTCAGCTGCGCGGACGCGGAGAAACCTCTTCAAGAAG 1620  
Db 2123 TGTGACAGCGGGGACTACAAGCTCAGCTGCGCGGACGCGGAGAAACCTCTTCAAGAAG 2182  
QY 1621 AAGTACAAGGCGCAGTATGTCCGCACTGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680  
Db 2183 AAGTACAAGGCGCAGTATGTCCGCACTGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242  
QY 1681 GCGAGGTTGATACGATGAGGCTGGGTGATGCGCGCCAGCGCCGAAACCTTCAAGCGG 1740  
Db 2243 GCGAGGTTGATACGATGAGGCTGGGTGATGCGCGCCAGCGCCGAAACCTTCAAGCGG 2302  
QY 1741 CACTGGCCAGAGGCGCCCTGAGGACCAAGATGCAAGATGGTGGGACTTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGAGGCGCCCTGAGGACCAAGATGCAAGATGGTGGGACTTTCAGTGGCACT 2362  
QY 1801 GGAGGCTTCCGACTACTAGCGCGCCAAACCCCAATTAAGTGAACATCGGTGTATCATC 1860  
Db 2363 GGAGGCTTCCGACTACTAGCGCGCCAAACCCCAATTAAGTGAACATCGGTGTATCATC 2422  
QY 1861 CTAGAGACGACACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGAGGCTGGAAA 1920  
Db 2423 CTAGAGACGACACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGAGGCTGGAAA 2482  
QY 1921 GACCAAGCTGCAATCGACACGAGATTGAACCCCTGCAAGAACAAAAATTAAGAACCTG 1980  
Db 2483 GACCAAGCTGCAATCGACACGAGATTGAACCCCTGCAAGAACAAAAATTAAGAACCTG 2542  
QY 1981 AGGGAAGTCCGAGTCACTGAAGAAAGCGCCAGAGAAATGAGTGTCTACAAATC 2040  
Db 2543 AGGGAAGTCCGAGTCACTGAAGAAAGCGCCAGAGAAATGAGTGTCTACAAATC 2602  
QY 2041 AGCTACCAACACCCAGCAAAAGGCGCTCAAGCAGACAGAGCTCCAGTCTGCACTCTTC 2100  
Db 2603 AGCTACCAACACCCAGCAAAAGGCGCTCAAGCAGACAGAGCTCCAGTCTGCACTCTTC 2662  
QY 2101 AGGAAGGCGCTGCAAGAGAGAGCAAGGTGTGTGTGGTGAGCAGAGCGCAGAGAG 2160  
Db 2663 AGGAAGGCGCTGCAAGAGAGAGCAAGGTGTGTGTGGTGAGCAGAGCGCAGAGAG 2722  
QY 2161 AAACCTCCGCAAGCTGCTCAAGCGCTGCAAGAACAGCAGCTCAGCTGCGAGGCTC 2220  
Db 2723 AAACCTCCGCAAGCTGCTCAAGCGCTGCAAGAACAGCAGCTGCGAGCTGCGAGGCTC 2782  
QY 2221 AGTGCTTCAACCCAGCAACACGACCTGGCAGACGCGCTTTCTTGGACACTGGGCGCT 2280  
Db 2783 AGTGCTTCAACCCAGCAACACGACCTGGCAGACGCGCTTTCTTGGACACTGGGCGCT 2842  
QY 2281 TTCTGTGCTGCACCGCGCCCAACATTAACAGTACTGTGTGATGAGGACCATCAATGAG 2340

Db 2843 TTCTGTGCTGCACCGCGCCCAACATTAACAGTACTGTGTCATGAGGACCATCAATGAG 2902  
QY 2341 ACTCACAATTTCTCTTCTGTGAATTTGCAATTTGCAATTTGCTTCTAGAGTACTTTGATCTCAAC 2400  
Db 2903 ACTCACAATTTCTCTTCTGTGAATTTGCAATTTGCAATTTGCTTCTAGAGTACTTTGATCTCAAC 2962  
QY 2401 ACAGACCCCTACCAAGCTGATGAATGCAAGTGAACACACTGGACAGGATGTCTCTCAACAG 2460  
Db 2963 ACAGACCCCTACCAAGCTGATGAATGCAAGTGAACACACTGGACAGGATGTCTCTCAACAG 3022  
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGAGTGCAGAGGTTTACAGCGAGTGAACCCCGG 2520  
Db 3023 CTACAGCTACAGCTCATGAGCTGAGGAGTGCAGAGGTTTACAGCGAGTGAACCCCGG 3082  
QY 2521 ACTCGAAACATGGAACCTGCGGACTTAAAGATGAGGAGAACTATGAGCAATACAGGCACTTT 2580  
Db 3083 ACTCGAAACATGGAACCTGCGGACTTAAAGATGAGGAGAACTATGAGCAATACAGGCACTTT 3133  
QY 2581 CAGCGTGAAGTGGCCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Db 3134 CAGCGTGAAGTGGCCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193  
QY 2641 GAAGCTGGGAAGGTTAAAGAAACACAGAGGTGAGACCTTCCAAACATAGAGGCACTACC 2700  
Db 3194 GAAGCTGGGAAGGTTAAAGAAACACAGAGGTGAGACCTTCCAAACATAGAGGCACTACC 3253  
QY 2701 TGACTGCAAGGCAATGAATAAACCATGTGGGTGATTTCCAGGACCTGTGCTATTGGCC 2760  
Db 3254 TGACTGCAAGGCAATGAATAAACCATGTGGGTGATTTCCAGGACCTGTGCTATTGGCC 3313  
QY 2761 AGGAGGCTGAGAAAGCAAGCAGCAGCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTGAGAAAGCAAGCAGCAGCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAAATCTCAGGAAGTCCATTTTCCCTGCTGCTTTGCTTGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATAAATCTCAGGAAGTCCATTTTCCCTGCTGCTTTGCTTGGATTA 3433  
QY 2881 TACCTCACAGCTGCACAAATGCAATTTTTCGTATCAAAAGTCAACCACTAACCCCTCCC 2940  
Db 3434 TACCTCACAGCTGCACAAATGCAATTTTTCGTATCAAAAGTCAACCACTAACCCCTCCC 3493  
QY 2941 CCAGAAAGCTCACAAGAGGAAACCGAGAGAGCGAGCAGAGAGATTTCTTGGAAATTTCT 3000  
Db 3494 CCAGAAAGCTCACAAGAGGAAACCGAGAGAGCGAGCAGAGAGATTTCTTGGAAATTTCT 3553  
QY 3001 CCCAAGGCGAAAGTCATTTGGAATTTTAAATCATAGGGAAGAGCAGTCTGTTCTAAA 3060  
Db 3554 CCCAAGGCGAAAGTCATTTGGAATTTTAAATCATAGGGAAGAGCAGTCTGTTCTAAA 3613  
QY 3061 TCCTCTTATTTCTTTGTTTGTCAAAAGAGGAACTAAGAGAGGAGCAGAGGCAAGCT 3120  
Db 3614 TCCTCTTATTTCTTTGTTTGTCAAAAGAGGAACTAAGAGAGGAGCAGAGGCAAGCT 3673  
QY 3121 GGAGGCTGAAAAACAGTGCAGAGACGTTTGAATGATGAGTCAAGTCAAGAGAGATGA 3180  
Db 3674 GGAGGCTGAAAAACAGTGCAGAGACGTTTGAATGATGAGTCAAGTCAAGAGAGATGA 3733  
QY 3181 CATTTACCTAGCACTATAAACCCTGTTGCTCTGAAGAAATGCTGCTTCTATTATATAT 3240  
Db 3734 CATTTACCTAGCACTATAAACCCTGTTGCTCTGAAGAAATGCTGCTTCTATTATATAT 3793  
QY 3241 GTGACTATTATCATGTAATCAACATGGGAACTTTTGGGGAACCTTAATTAAGAAATCCCAA 3300  
Db 3794 GTGACTATTATCATGTAATCAACATGGGAACTTTTGGGGAACCTTAATTAAGAAATCCCAA 3853  
QY 3301 TTTTCAGGAGTGGTGGTCAATAAACCTCTGTGGCCAGTGTAAAAAGAAAA 3353  
Db 3854 TTTTCAGGAGTGGTGGTCAATAAACCTCTGTGGCCAGTGTAAAAAGAAAA 3906

Search completed: February 14, 2004, 14:35:40



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:35:48 ; Search time 221 Seconds  
(without alignments)  
6716.613 Million cell updates/sec

Title: US-10-025-966A-4

Perfect score: 3363

Sequence: 1 gactccgcgcatcccaaga.....aaagaaaaaaaaaaaaaa 3363

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCUTS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3323.4	98.8	3906	4	US-09-996-243-83
2	1308.6	38.9	3057	4	US-09-886-319A-53
3	1073	31.9	5769	4	US-09-668-673B-1
4	1040.4	30.9	2616	4	US-09-773-426A-12
5	1040.4	30.9	4321	4	US-09-773-426A-4
6	1002.4	29.8	4834	4	US-09-668-673B-15
7	1001.6	29.8	1367	4	US-09-668-673B-13
8	578.2	17.2	1799	4	US-09-810-347-1
9	441	13.1	441	4	US-09-404-879A-315
10	381.6	11.3	494	4	US-09-668-673B-10
11	375.8	11.2	466	4	US-09-668-673B-9
12	369	11.0	3871	4	US-09-484-970B-70
13	368	10.9	459	4	US-09-668-673B-12
14	347.4	10.3	538	4	US-09-668-673B-8
15	297.4	8.8	436	4	US-09-668-673B-11
16	291	8.7	296	4	US-09-668-673B-21
17	288.4	8.6	302	4	US-09-313-294A-7135
18	210.8	6.3	540	4	US-09-668-673B-5
19	177	5.3	482	4	US-09-668-673B-6
20	172	5.1	3577	4	US-09-620-312D-457
21	157.4	4.7	288	4	US-09-668-673B-19
22	156.6	4.7	42571	4	US-09-810-347-3
23	104	3.1	590	4	US-09-668-673B-17
24	99.4	3.0	614	4	US-09-668-751-22
25	74	2.2	473	4	US-09-668-673B-4
26	59.2	1.8	7218	1	US-08-232-463-14
27	50.4	1.5	1859	3	US-08-894-818B-15

28	50.4	1.5	1977	3	US-08-894-818B-2	Sequence 2, Appli
29	50.4	1.5	1977	3	US-08-894-818B-6	Sequence 6, Appli
30	50.4	1.5	1977	4	US-09-445-472-11	Sequence 11, Appl
31	49.4	1.5	505	4	US-08-961-527-389	Sequence 389, App
32	49.4	1.5	5552	3	US-08-153-888-1	Sequence 1, Appli
33	48.8	1.5	290	4	US-09-444-791A-113	Sequence 113, App
34	48.8	1.5	1926	4	US-09-249-585A-2	Sequence 2, Appli
35	48.8	1.5	1926	4	US-09-410-399-3	Sequence 3, Appli
36	48.8	1.5	2580	3	US-09-050-863-2	Sequence 2, Appli
37	48.8	1.5	2580	4	US-09-359-081-2	Sequence 2, Appli
38	48.8	1.5	5452	2	US-09-130-114-1	Sequence 1, Appli
39	48.8	1.5	8705	4	US-09-647-344A-14	Sequence 14, Appl
40	48.8	1.5	9600	3	US-08-910-647-1	Sequence 1, Appli
41	48.8	1.5	9600	4	US-09-620-925-1	Sequence 15, Appl
42	48.8	1.5	10596	1	US-07-884-811-15	Sequence 15, Appl
43	48.8	1.5	10596	1	US-07-885-971-15	Sequence 15, Appl
44	48.8	1.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
45	48.8	1.5	10596	1	US-08-194-088B-15	Sequence 15, Appl

## ALIGNMENTS

### RESULT 1

US-09-996-243-83  
; Sequence 83, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desrochers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28









QY	1301	-TGCTACACAAGAGAGACAATGACAAGGTGGACGCCCGAGGAGAGAACTTTCTGCCCAG	1359
Db	362	TTGCTCCACAAGAGGAGGGTGACAAAGTGAATGCCCCAGGAGGAACTTCTTGCCCCAG	421
QY	1360	TACACGGCTGAAGGACCTGTGTACAGCTGCTGAGTACCAGACGGCGTGTGAGCAGCTG	1419
Db	422	TACCACGGCTGAAGACCTGTGTACAGCAGCTGAGTACCGACACAGCATCGGAACAGCTG	481
QY	1420	GGACAGAATGGCAGTGTGTGGAGGACGCCACGGGGAAAGCTGAAGCTGTCATAAGTGCAG	1479
Db	482	GGGCAGAAATGGCAGTGTGTGGAGGACCGCTTCTGGGACCGCTGAAGCTGCAACAATGTAA	541
QY	1480	GGCCCCATCGGGCT-----GGGGCGGACAGACAGCCCTCTCCAACTCTGTCGCCAG	1530
Db	542	GGCCCCATCGGCTTGGTGGCGGGCTGGACAGCAGAGCCCTCTCCAACTCTGTCGCCAG	601
QY	1531	TACTACGGGCAGGGCAGGAGGCTGCACCTGTGACA-----GCGGGGACTACAAG	1581
Db	602	TATGACGGGCAGACAGCAGGAGGCTGCAGCTGTGAAGTGGCGGTGGAGGGGACTACAA	661
QY	1582	CTCAGCTGCGCGGACCGCGGAAAACTCTTCAAGAAGAAGTACAAAGGCCAGCTATGTC	1641
Db	662	CTGGGCTGCTGGACGCCG---TAAGCTCTTTAAGAAAGAATATGAACCAAGCTATGCC	718
QY	1642	CGCAGTCTGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGGAGGTGTACCAAGCTAGGC	1701
Db	719	CGAAACCGCTCCATCCGTTCCGTGGCCATCGAGGTGGACGGTGAATATACCAAGCTAGGC	778
QY	1702	CTGGGTGATCGCGCCAGCGCCGAAACCTCACCAAGGGCACTGGCCAGGGGCCCTGAG	1761
Db	779	TTGGATCTGTCCTCAGCCCGCAACCTTAGCAAGCCGCACTGGCCAGGGGCCCTGAA	838
QY	1762	GACCAAGATGACAAGGATGCTGGGGACTTCAGTGGCACCTGAGGCGCTTCCCGGACTACTCA	1821
Db	839	GACCAAGATGACAAGGATGCTGGCAGTTTCAGTGGTACTCGTGGCTTCCAGATTATCT	898
QY	1822	SGCGCCAAACCCATTAAAGTGACATCGTGTCTATCTCTAGAGAAGCAAGTCCAG	1881
Db	899	GCCCCAAATCCCATCAAAGTGAACCATCGGTGTATCATCTTGAGAAATGACACAGTCCAG	958
QY	1882	TGTGACTGGACCTGTCAAGTCCCTGCGGCGCTGGAAAGACCAAGCTGACATCGAC	1941
Db	959	TGCGACTTGGACCTGTACAGTCCCTGCGGCTTGGAAAGACCACAAGCTGCACATCGAC	1018
QY	1942	CACGAGATTGAACCTTGAGAACAAATTTAAGAACTTGAGGGAAATCCGAGTCACTTG	2001
Db	1019	CATGAGATCGAAACCCCTGCGAACAATAATTAAGAACCTTCGAGAAGTCAAGGGTCACTTG	1078
QY	2002	AAGAAAAGCGGCGCAGAAAGATGTGACTGTCAAAAATCAGCTTACCAACCCAGCACAA	2061
Db	1079	AAGAAGAGCGACCGGNAAGATGTGACTGCCATATATCAGTTACCAAGCCAAACAA	1138
QY	2062	GGCGGCTCTAAGCACAGAGGCTCCAGTCTGCTATCTTTTCAAGGAAGGCGCTGCAAGAAG	2121
Db	1139	GGCGGCTCTAAGCACAAAGGCTCCAGCTGCACCCCTTTCAGGAAGGGTCTGCGAGGAAG	1198
QY	2122	GACAAGGTGTGGCTTGTTCGGGAGCAGAGCGCAAGAAAGAAATCCGCAAGCTGCTCAAG	2181
Db	1199	GACAAGGTGTGGCTGTGCGGGA-CAGAAACGCAAGNAGAAATGCGCA-CTGCTCAA	1256
QY	2182	CGCGTGCAGAAACAGCACAGTGAGCATGCCAGGCTCAAGTGTTCACCCACGACAAC	2241
Db	1257	CGCGTGCAGAAACAGCATGTCGAGCATGCCCGGCTTCAAGTGTTCACCCAGCAAC	1316
QY	2242	CAGCACTGGCAGACGGGCGCTTTCTGAGCACTGGGGCTTTTCTGTGCTTGCACACAGCGCC	2301
Db	1317	CACCATGGCGAGACGGGCCCACTCTGACGCTGGGGCGGTTCTTGCGCTTGCACAGCGCC	1376
QY	2302	AACAATAACAGTACTGGTGCATGAGGACCAATCAATGAGACTCAAAATTCCTTCTGT	2361
Db	1377	AACAAACAAACGTAAGTGGTGTGAGGACCAATAATGAGACCCACAACTTCTTCTGTG	1436

```

RESULT 3
US-09-668-673B-1
; Sequence 1, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dnoco, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5769
; TYPE: DNA
; ORGANISM: quails
US-09-668-673B-1

```

Query Match	31.9%	Score 1073	DB 4	Length 5769
Best Local Similarity	65.9%	Pred. No. 1.3e-265		
Matches 1664: Conservative	4	Mismatches	809	Gaps 6
		Indels	48	

Qy	97	CTGCTGGTGGAAAGCTCGGCCTTCTGTGCGACCAACCGCTGAAAGCGCAGGTTCAGAGG	156
Db	788	CTGCTGACAACTCATTTCTCCACTCTCAAAGTCCTGAGGTTACAGAGCGCGTGTCAGCAA	847
Qy	157	GACCGCAGGAACATCGCCCCAACATCATCTCTGTGCTGACGACAGCACCAGGATGTGGAG	216
Db	848	GAGAGAAAATATCAGACCAATATCATCTCTGTGCTACAGATGACCAAGATGTGGAG	907
Qy	217	CTGGGTTCATGCAAGTGATGAAACAAGACCGCGCGCATCATGGACAGAGCGCGGGCGCAC	276
Db	908	CTAGGTTCCTTCAAGTGATGAACAAACACGACCGATTATCGAGAATGAGGGGCATCC	967
Qy	277	TTCATCAAGCCTTCTGTGACCACACCCATGTGTGTCGCCCTCAGCTCTCTCCATCCTCACT	336
Db	968	TTCATCAATGCCCTTCGTATAACACCCCGATGTGTGTCGCCATCAGTTCCTCCATGCTGACT	1027
Qy	337	GGCAAGTAGTGTCCCAACACACACACTCACACCAAGATGAGAACTGCTCCTCGCGCTCC	396

1028	CGAAGTATGTCACCAACCAACACATACACCAACATGAACCTCTCTCCCTCC	1087	2108	CCAGGACAGAGTGGCAGTGACAGAGATGCTTCTGGCAAGCTTCGAATTCACAAGTGC	2167
397	TGGCAGGACAGACGAGAGCGGACCTTTGGCGGTGTACCTCAATAGACGTGGTACCGG	456	1477	AAAGGCCCCA-----TGGGCTGGCGGACGAGAGCCCTCTCCAACTCTGGTCCCAAGTAC	1533
1088	TGGCAGGCGACCTCAGCAGCCAGCGACTTTCCGCGGTGTATCTGAATAACATCTTGGGTATCGA	1147	2168	AGGTTATCTAGTGACATCTCTGCGCATCAGGAAAGGACCCGCGCATCACTTCCAGGGGA	2227
457	ACAGCTTTCTTCCGGAAGTATCTTAATCAATACACAGGCTCTCTAGTGCACCCGGCTGG	516	1534	TACGGCAGGCGAGGAGGCTGTGACAGCGGGACTACAAGCTCAGCTCAGCTTGCC	1593
1148	ACAGCTTTTGGGAATACCTCAATGATACATGCGCAGCTACCTCTCTGGGTGG	1207	2228	TAVAGTGGTAAAGATAAGGACTGCACTGTGGAGACACCGATTTCCGAAAACAGCAGGACC	2287
517	AAGGAGTGGTGGGACTCTTTAAAACTCCGCTTTTATACATACAGCTGTGTGGAAC	576	1594	GGAGC-----CGGAAAAAATCTTCAAGAAAGTACAGGCCAGC	1635
1208	AGAGAGTGGGTGGATTAAGTGAAGAACTCTGCTTCTATAATACACCATTTCTCGCAAT	1287	2288	CAAGAAAAAATCAAGGAGGTTCTGAGAAACCCAGTGGCGCAAAAAATACAAACACAT	2347
577	GGGTGAAAGAAAGCAGCGCTCCGACTACTCCAGGATTAACCTCACAGCTCATCACC	636	1636	TATGTCGCGAGTCGCTCCATCCGCTCAGTGGCCATCAGAGGTGACCGGAGGTGTACAC	1695
1268	GGTAACAAAGAGAGCATGGAATGATATGGAAGGACTACTTCAAGACTTAATCACT	1327	2348	TTTGTTCACACTCGCCAAACCCGGTCTTGTGAGTGAATTTGAAGGTGAATAATATGAC	2407
637	AATGACAGCTGAGCTTTCTCGACGCTCCAGAGATGTACCGGACAGCGGCTCTC	696	1696	GTAGGCTGGGTGATCCGCCC-----AGCCCCGAACTCACCAGGCGGAC	1743
1328	AATGAGAGCATTAATTAATCTTCAAGATGTCCAGAGGATATACCCATAGGCCCATATG	1387	2408	ATAAACCTGGAAGAGGAAGAACTGAGGTGTTAAAGACGAGAAATATACCAAAAGTAC	2467
697	ATGCTCATCAGCATGACGCCCCCAGCGCTGAGGATTCAGCCCCACATATTCACGC	756	1744	TGCGCAGGGGCCCTGAGGACCAAGATGACAAAGGATGG-----TGGGACTTCAAGTGC	1797
1388	ATGGTCATCAGCCATGCTGGGCTCATGGCCCTGAGGATTCGGCCCCACAGTTCTCAG	1447	2468	AATGCTGAAAAATGACAAAAAAGCAGAGGAACTGATGTCTCTGTGACACGATGTT	2527
757	CTCTTCCCAAGCGCATCTCAGCACATCAGCGGAGCTACAACTAGCGGCCCAACCCGAC	816	1798	ACTGAGGCTTCCCGACTACTCAGCGCCAAACCCATTAAGTGACACATCGGTGCTAC	1857
1448	CTTACCCCAACGCTTCAAGCATATACCCGCCACTATACTATGACCAACAAATGAT	1507	2528	GCTGATGGCACTGATTTATAGGTCAACCCAGTTCTGTGAGAGTGACACAAAGTGT	2587
817	AAACACTGATCATGGCTTACAGGGGGCCATGAAGCCATCCACATGGAATTCACCAAC	876	1858	ATCTAGAGAAACGACACAGTCCAGTGTGACCTGACCTGTACAAGTCCCTGCGGCGCTGG	1917
1508	AAGCACTGATCATGAGTACACGGGGCCCATGTGCTCTCCATGCAATGGAGTTTACAAAC	1567	2588	ATTCTTCAATGACACTATTCCTGTGAGAGGAGCTGTACCAATCTGCCAGAGCTGG	2647
877	ATGCTCCAGCGAAGCGCTTGCAGACCTCTCATGCTGGTGGAGCTTCCATGGAGACGATT	936	1918	AAAGACCAAGCTGCATCCGACCAAGATTTGAAACCTCTCGACAAACAAATTAAGAAC	1977
1568	GTCTTGCACGCAAGACTTCAGACCTGTATGTCAGTTGATGACTCTATGGAAGATTA	1627	2648	AGGACCAACAGGCTTACATCGATAAGGAGATTGAAGCTCTCCAGGACAAAATCAAGAT	2707
937	TACACATGCTGGTGAAGCGGCGAGCTGGAACAACATGATATACACCGCGAC	996	1978	CTGAGGGAAGTCCGAGTCCCTGAGAAAAAGCGCCAGAGAAATGTGACTGTCAAAA	2037
1628	TACCAATGCTTGCAGAAATGGGAGAGCTGGAGAAATACCTCATTAATTTACACAGCTGAC	1687	2708	TTGAGGGAAGTTAGAGGACACCTTAAGAGAAACACAGACGATGTGACTGTACTAAA	2767
997	CAGGTTACCATCGGCCAGTTTGGCTGTGGAAGGAAATCCATCCCATATGATGTT	1056	2038	ATCAGTACCAACCCAGCACAAAGCC-----CGCTCAAGCAGAGGCTTCCAGTCTG	2091
1688	CATGTTACCATATGGCGAGTTTGAATGCTGTCAGGAGGAGTCAATGCCATATGACTTT	1747	2768	CAGAGCTACTACAAAGAGAGAAAGCGGTAAAGACCAAGAGAAAAATCAAGAGCATCTA	2827
1057	GACATCAGGCTCCGTTCTACGTGAGGGGCCCAACAGTGAAGCGGCTGTCTGAATCCC	1116	2092	CATCTTTTCAAGNAGGCGCTGCAAGAGGA-----CAAGGTGTGCTGTGGCGGAGCAG	2148
1748	GATATTCGAGTTCTTCTTTTATTCGTTCCAAAGTGTAGAGCCAGGATCTGTAGTGCT	1807	2828	CATCCCTTCAAGAGAGCAGCAGGAGGTAGACAGCAAACTGCAGCTGTTCAGAGAGAT	2887
1117	CACATGCTCTCAATGACCTGGCCCCCACCCTCTGGAATTCAGGCTGAGGCTGGACATA	1176	2149	AAGCGCAAGAGAAACTCCGCAAGCTGCTCAAGCGCTTCAGAAACAAACGACAGCTGCAGC	2208
1808	CAGATAGTTCTGAATATGATCTTGACCAACAAATCTTGATATTCAGGACTTGACACA	1867	2888	CGCAGAGAGAGAGAGAAAGGAAAGAAAGCGCCAGAAAGGCGGATGAGTGTAGC	2947
1177	CTTCCGATATGACGGGAAATCCATCTCTCAAGCTGCTGGAACGAGCGGCGGTGAAT	1236	2209	ATGCCAGGCTCAGTGTCTTACCACGACAAACGAGACTGCGCAGACGCGCTTCTCG	2268
1868	CTTCCAGATATGATGCGCAATCTGTCTTAAAGCTTCTAGACTTGGAGAGACAGGAAT	1927	2948	CTTCTGAGCTGACATGTTTACTCATGACAAATAACATTTGGCAAACTGCACCTTCTCG	3007
1237	CGGTTTCACTTGAAGAAAGATGAGGCTCTGGCGGACTCTCTTCTGGTGGAGAGGC	1296	2269	ACACTGGGCTTTCTGTGCTGCGCAGCGGCAACAAATAACACGCTACTGTGTGCTAGAG	2328
1928	AGGTTTCGAACAAACAGAGAGCCAAATCTGGCGTGACACATCTCTGGTGAAGAGGC	1987	3008	AACTTGGGATCTTCTGTGCTGTGCAAGCTCAATATCAACACTTACTGTGTGTGCGA	3067
1297	AAGCTGCTACAAAGAGACATGACAGGTGGAGCGCCGAGGAGGAATCTTCTGCCC	1356	2329	ACCATCAATGAGACTCAATTTCTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTAC	2388
1988	AAATTTCTGGCAAAAAGAGGAGCTACAAAACACTTCAGCAATCTAATCAACTGCCA	2047	3068	ACAGTGAATGACACCCCAATTTCTCTTGTGTGATTTGCAACTGGCTTCTTGGAAATWC	3127
1357	AAGTACAGGCTGTGAAGACCTGTGTAGCGTGTGAGTACACAGCGGCTGTGAGCAG	1416	2389	TTTGTCTCAACACAGACCCCTTACAGCTGATGATGCTGAGTGAACACACTGGACAGGAT	2448
2048	AAGTATGAGAGGTTAAAGAAATTTATGCCAACAGCGAGATACACAGACGCTGTGAAACA	2107	3128	TTTGTATGATGACTGACCCCTTATGAGTGTGACAAATACCGTACATACAGTGGAAAGGCG	3187
1417	CTGGGACAGAGTGGCAGTGTGTGGAGGAGCGCCACGGGGAAGCTGACATAGTGC	1476	2449	GTCTCAACAGCTACACGCTACAGCTCATGGAGCTGAGAGCTGCAAGGCTTACAGCAG	2508
			3188	ATTTTAAATCAATTACATGTACATGTTAATGGAATTAACGAAGTTGTCAAGGTTATAGCAG	3247

QY 2509 TGTAACCCCGGACTCGAAACATGAGCTGGACCTTAAGAGTGAAGGAGCTATGAGCAA 2568  
DB 3248 TGCAATCCGAGCGCGAAGGACTTGAACAGGAATAAAGATGAGGAGCTATGATCCA 3307  
QY 2569 TACAG 2573  
DB 3308 CACAG 3312

RESULT 4  
US-09-773-426A-12  
; Sequence 12, Application US/09773426A  
; Patent No. 6534302  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria Alexandra  
; APPLICANT: Williams, Mark  
; APPLICANT: Tsia, Fong-Ying  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302e1  
; FILE OF INVENTION: Human Sulfatases (A CIP Application)  
; FILE REFERENCE: 35800/208398(5800-79  
; CURRENT APPLICATION NUMBER: US/09/773,426A  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/495,823  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 2616  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-773-426A-12

Query Match 30.9%; Score 1040.4; DB 4; Length 2616;  
Best Local Similarity 65.5%; Pred. No. 2.2e-257;  
Matches 1660; Conservative 0; Mismatches 816; Indels 60; Gaps 7;

QY 98 TGCTGGTGGAGCTGGCTTCCTGTCGCCACCCGCTGAAGGAGGAGGTTTCAGAGG 157  
DB 50 TGCTGGAGGCTCTGTTGGACTGTGAGATCCCGAGTCCGAGGTCAGAGGAGGATACAGCAGG 109  
QY 158 ACCGAGGAAACATCCGCCCCCAACATCATCTGCTGCTGACGACGACGACGAGGATGAGG 217  
DB 110 AACGAAAACATCCGACCCACATTTCTGTGCTTTACCGATGATCAAGATGAGG 169  
QY 218 TGGGTTCCATGAGGTGATGAACAAGACCGCGGATCATGAGAGCAGGCGGGGCGACT 277  
DB 170 TGGGTTCCCTGCAAGTCTATGAACAAACGAGAAAGATTATGGAACATGCGGGGGCCACT 229  
QY 278 TCATCAAGCCTTCGTGACCAACCCATGCTGCTGCCCTCAGCTCCCTCCATCTCACTG 337  
DB 230 TCATCAAGCCTTCGTGACCAACCCATGCTGCTGCCCTCAGCTCCCTCCATGCTCACTG 289  
QY 338 GCAAGTACGTCACCAACCAACACCTACCAACAAATGAGAACTGCTTCCTCGCCCTCT 397  
DB 290 GGAAGTATGTCACAATCAATGCTCTACCAACAAACGAGAACTGCTTCCTCCCTCGT 349  
QY 398 GCGAGCAGCAGCAGAGGCGCCGCTTTCGCTGCTACCTCAATAGCACTGCTACCGGA 457  
DB 350 GCGAGCCTATGATGAGCTCGGCTTTGCTGTATATCTTAAACAACTGGCTACAGAA 409  
QY 458 CAGCTTTCTTCGGGAAGTATCTTAATGAATCAACGGCTCTCAGTGGCCACCCGCGTGA 517  
DB 410 CAGCTTTCTTCGGGAAGTATCTTAATGAATCAACGGCTCTCAGTGGCCACCCGCGTGA 469  
QY 518 AGAGTGGCTCGGACTCTTAAAACTCCGCTTTTATTAATCAACGCTGTGTCGAAG 577  
DB 470 GAGAATGGCTTGAATTAATCAAGAAATTCGCTTCTTAAATTAACACTGTTTGTGCAATG 529  
QY 578 GGGTGAAGAAAGACGCGCTCCGACTCTCCAGAGATTTACCTCAAGACTCAGACTCATCA 637  
DB 530 GCATCAAGAAAGCATGGAATTTGATATGCAAGGACTACTTCAAGACTTAATCACTA 589

QY 638 ATGACAGCGTGGCTTCTTCCGACGCTCCAGAAAGATGTACCGGCACAGCGCCAGTCTCA 697  
DB 590 ACAGAGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 649  
QY 698 TGGTCAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 757  
DB 650 TGGTCAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 709  
QY 758 TCTTCCAAACGATCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 817  
DB 710 TGTACCCCAATGCTTCCCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 769  
QY 818 AACACTGGATCATGCGCTACACGCGGCGCCCATGAAGCCCATCCACATGGAATTCACAA 877  
DB 770 AACACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 829  
QY 878 TGCTCCAGCGGAGCGCTTGCAGACCCCTCATCTCGGTGGAGCAGCTCCATGGAGAGGATTT 937  
DB 830 TTCTACAGCGCAAAAGGCTCCAGACTTTTGTGTGATGATGATGATGATGATGATGATGAT 889  
QY 938 ACAACATGCTGCTGAGACGCGGCGAGCTGGACACACACGCTACATCGTATACACCGCG 997  
DB 890 ATAACATGCTGCTGAGACGCGGCGAGCTGGAGATATTTACATCATTTACACCGCGACC 949  
QY 998 ACGGTTACCAATCGCCAGTTTGGCTGCTGGAAGGAAATCCATGCCATATGATTTG 1057  
DB 950 ATGGTTACCATATGGCGAGTTTGGACTGCTGGAAGGAAATCCATGCCATATGATTTG 1009  
QY 1058 ACATCAGGCTCCGCTTCTACGTGAGGCGGCCAACTGGAAGCGGCTGTCTGAAATCCCC 1117  
DB 1010 ATATTCGTGTGCTTTTATTTATTTGCTGCTCAAGGTAGACCAAGATCAATAGTCCCAC 1069  
QY 1118 ACATGCTCTCAACATGAGCTGGGCCCCCACCATCTCTGAGCATTTGAGCGCTGGACATAC 1177  
DB 1070 AGATGCTTCTCAACATGAGCTGGGCCCCCACCATCTCTGAGCATTTGCTGGCTCGACAC 1129  
QY 1178 CTGCGGATATGAGCGGAAATCCATCTCAAGCTGCTGACACGAGCGCGCGGCTGAATC 1237  
DB 1130 CTCTGATGTGAGCGCAAGTCTGCTCTCAAACTTCTGACCCAGAAAGCCAGGTAAAC 1189  
QY 1238 GGTTCACATGAAAAAGATGAGGCTCTGCGGAGCTCTCTTCTTGTGGAGAGGCA 1297  
DB 1190 GGTTCGAAACAAAGAGGCGCAAAATTTGGCGTGATACATTTCTTAGTGGAAAGAGGCA 1249  
QY 1298 AGCTGTACACAGAGACAAATGACAGGTGAGCGCCAGGAGGAGACTTTCTGCCCA 1357  
DB 1250 AATTTCTACGTAAAGAGGAAATCCAGCAAGATATCCAAAGTCAATCACTTGGCCCA 1309  
QY 1358 AGTACAGCGGTGTAAGGAGCTGTGTGACGCTGCTGAGTACAGAGCGGCTGTGAGCAGC 1417  
DB 1310 AATATGAACGGGTCAAGAACTATGCCAGCAGGCGCAGGTACAGAGAGCTGTGAACAAC 1369  
QY 1418 TGGGACAGAGTGGCAGTGTGTGAGGAGCGCCACGCGGAGCTGAGCTGCAATAGTGA 1477  
DB 1370 CCGGGCAGAGTGGCAATGCAATGAGGATACATCTGCGCAAGCTTCGAATTCAGAGTGA 1429  
QY 1478 AGGCGCCCATGCGCTGGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1525  
DB 1430 AAGGACCCAGTACCTGCTCAGTCCGCGCAGAGCAGCGCGGAACTCTACGCTCGCGCT 1489  
QY 1526 CCAAGTACTACGCGCAGG-----GCAGCGGCTGCGACCTGTGTGACAGCGGGA 1576  
DB 1490 TCCATGACAAAGACAAAGAGTGCAGTTGTAGGAGTCTGTTACCGTGCAGCAGAGGCG 1549  
QY 1577 ACAAGTCTGCGCTGGCGCGGCGGAGGAAACTCTTCAAGAGAGTACAGAGGCGGCGCT 1636  
DB 1550 AAGAAAGAGTCAACGGCAATTTTGAAGAAACAGGGGACTCCAAAGTACAGAGGCGGCT 1609  
QY 1637 ATGTCGCGCAGTCCCTCCATCCGCTCAGTGGCCATCGAGGTGAGCGGAGGCTGTACACG 1696  
DB 1610 TTGTCATACTCGGCGAGACACGTTCTCTTCTGCGTGAAATTTGAAGGTGAAATATATGACA 1669

1697 TAGGCTGGGTGATCGG-----CCAGCCCGGAAACCTCCACCAAGCGC 1741  
1670 TAAATCTGGAGAGAGAGATTCCAGTGTTCACCAAGAAACATTTGTAAGCGTC 1729  
1742 ACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGGATGGTGGGACTTTCAGTGCGACTG 1801  
1730 ATGATGAAGGCCCAAGAGGCCAAGAGATCTCCAGGCTTCCAGTGGTGCAACAGGGCA 1789  
1802 GAGGCTTCCGACTACTCAGCGCC-----NACCCCATTAAGTGACACATC 1849  
1790 GGATGCTGGCAGATAGCAGCAAGCGCGTGGGCCACCTACCACTGTCCGAGTGACACACA 1849  
1850 GGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGCG 1909  
1850 AGTGTTTTATCTTCCCAATGACTCTATCCATTGTGAGAGAGACTGTACCAATCGGCA 1909  
1910 AGGCTTGAAGACCAAGACTGACATCGACACGAGATTTGAACCTTGCAGAACAAAA 1969  
1910 GAGCGTGAAGGACCAATAGGCATACATTTGACAAAGAGATTGAAGCTTGCAGATATAA 1969  
1970 TTAAGAACCTTAGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGGCAGAAAGATGTGACT 2029  
1970 TTAAGAAATTAAGAGAGTGAAGGACATCTGAGAGAGAGAGCTGAGGATGTAGCT 2029  
2030 GTCACAAATACGTACCAACACCCAGCAAAAGGCCGCTCAAGC-----ACAGAGGCT 2083  
2030 GCAGTAAACAAAGCTATTACATAAAGAGAGAGAGGTGTAAAAAAGCAAGAGAAATTAAGA 2089  
2084 CCAGTCTGCACTCTTTCAGGAAG--GGCTGCAAGAGAGAGCAAGGTGTGGCTGTGTC 2140  
2090 GCCATCTTACCATTTCAGGAGGCTGCTCAGGAAGTAGATAGCAAACTGCAACTTTTCA 2149  
2141 GGAGCAGAGAGCGCAAGAAACTCCGCAAGTGTCTAAGCGCCTGCGAGAACAA---CG 2197  
2150 AGGAGAACACCGTAGSAGNAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2209  
2198 ACAGCTGACGATGCCAGGCTCAGCTGCTTACCCACGACAGCAGCTGCGAGAGCG 2257  
2210 AAGAGTGCAGCTGCTGCGCTCCTCCTGCTTACCGCATGACAAACCACTGGCAGACAG 2269  
2258 CGCTTTCTGGACACTGGGSCCTTTCTGCTGCTGACACAGCGCCCAACAATAACAGTACT 2317  
2270 CCCGTTCTGGAACCTTGGGATCTTTCTGCTGCTGACAGAGTTCTAACATAACACTACT 2329  
2318 GGTGCATGAGACCATCAATGAGACTCAAAATTTCTCTTCTGTAATTTGCAACTGCTGCT 2377  
2330 GGTGTTTTCGCTACAGTTAATGAGACGCAATTTTCTTCTGAGTTTGTCTACTGGCT 2389  
2378 TCCTAGAGTACTTTGATCTCAACACAGACCCCTACCAAGCTGTAGTAATGCAATGACACAC 2437  
2390 TTTTGGAGTATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2449  
2438 TGGACAGGAGTCTCTCAACAGCTACAGCTACAGCTATGAGAGTGAAGAGCTGCAAG 2497  
2450 TAGAACGAGGCAATTTGAAATCAGCTACAGCTACAGCTACAGCTACAGCTACAGCTACAG 2509  
2498 GTTACAGAGCTGTAAACCCCGGACTCGAAACATGACCTGGGACTTAAAGATGGAGAA 2557  
2510 GATATAAGCAGTGAACCCCAAGACCTAAGAAATTTGATGTTGGAATAAAGATGGAGAA 2569  
2558 GCTATGAGCAATACAG 2573  
2570 GCTATGACCTACACAG 2585

## RESULT 5

US-09-773-426A-4  
; Sequence 4, Application US/09773426A  
; Patent No. 6534302  
; GENERAL INFORMATION:  
; APPLICANT: Gluckman, Maria Alexandra  
; APPLICANT: Williamson, Mark  
; APPLICANT: Tsia, Fong-ying

; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el  
; FILE REFERENCE: 35800/208398(S900-79  
; CURRENT APPLICATION NUMBER: US/09/773,426A  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/495,823  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 4321  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (510)...(3125)  
; NAME/KEY: misc\_feature  
; LOCATION: 4310  
; OTHER INFORMATION: n = A,T,C or G  
US-09-773-426A-4

Query Match 30.9%; Score 1040.4; DB 4; Length 4321;  
Best Local Similarity 65.5%; Pred. No. 2.8e-257;  
Matches 1660; Conservative 0; Mismatches 816; Indels 60; Gaps 7;  
QY 98 TGTGCTGGTGAAGCTCGGCTTCTCTGTGCGCACCCCGCTGAAAGCGCAGGTTTCAGAGGG 157  
Db 559 TGTGGGAAGCCCTCTGTTCGACTGTGAGATCCCGAGGTTTCAGAGGACGATACACAGG 618  
QY 158 ACACGAGGAATCCGCCCCCAACATCATCTGCTGTGCTGCGGACGACGACGATGTGAGC 217  
Db 619 AACGAAAAACATCCGACCCCAACATTAATTTCTGTCTTACCGATGATCAAGATGTGAGC 678  
QY 218 TGGGTTCCATGCAAGTGTGATGAACAAGACCCCGGCGCATCATGAGAGCGGCGGCGCACT 277  
Db 679 TGGGTTCCCTGCAAGTCAATGAACAAACGAGAGAGATTATGGAACATGSGGGGCGCACCT 738  
QY 278 TCATCAAGCGCTTCTGTGACCAACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 337  
Db 739 TCATCAATGCTTGTGACTACACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 798  
QY 338 GCAAGTACGTCCCAACACACACACCTTACACCAACATGAGAACTGCTCTCTGCTGCTGCT 397  
Db 799 GGAAGTATGTGCAATCAATGCTTACACCAACAAACGAGAACTGCTCTTCTGCTGCTGCT 858  
QY 398 GGCAGGCAACGACGAGCGCGCACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457  
Db 859 GGCAGGCAATGCTAGCTGAGCTTGGACTTTTCTGTATATCTTAACAACACTGGGCTACAGAA 918  
QY 458 CAGCTTTCTTGGGAGTATCTTAATGAATACAAACGCTCTACGTGCCACCGGCTGGA 517  
Db 919 CAGCTTTTGGGAAATACCTCAATGAATATATGCGACTACATCCCTGCTGCTGCTGCTGCT 978  
QY 518 AGGAGTGGGTCGGAAGCTCTTAAAGACTCCGCTTTTATTAATCAAGCTGTGCGGAAG 577  
Db 979 GAGAAATGCTTGGATTAAATCAAGAAATCTCTGCTTCTATTAATTAATCTGCTGCTGCT 1038  
QY 578 GGTGAAAGAAAGACGCGCTCGACTTCCAGGATTCCTCAAGGATTCCTCAAGACTCATCA 637  
Db 1039 GCATCAAGAAAGACATGGAATTTGATTAATCAAGGACTACTTTCACAGACTTAATCACTA 1098  
QY 638 ATGACAGCGTGAAGTCTTCTGCGACGCTCCAAAGAGATGTACCGGCAACAGGCCAGTCTCA 697  
Db 1099 ACAGAGCATTAATTAATTTCAAAATGCTTAAGAGATGATATCCCATAGGCCGCTATGA 1158  
QY 698 TGGTCATCAGCCATGAGCGCCCGCCCAAGGCTTGAAGATTCAGGCCCAATATTCAGGCC 757  
Db 1159 TGGTCATCAGCCATGAGCGCCCGCCCGCCCGGAGACTCAGGCCCAAGTCTTCTTAAC 1218  
QY 758 TCTTCCCAAGCGCATCTCAGACATCAGCGGAGCTACAACTACGCGCCCAACCGGACA 817  
Db 1219 TGTACCCCAATGCTTCCCAACACATAACTCTTAGTTAATTAATGACCAATATGGATA 1278





ORGANISM: Homo sapiens  
US-09-668-673B-15

Query Match 29.8%; Score 1002.4; DB 4; Length 4834;  
Best Local Similarity 65.7%; Pred. No. 1.7e-247;  
Matches 1595; Conservative 0; Mismatches 771; Indels 60; Gaps 7;

QY 208 GATGTGGAGCTGGTTCCATGAGGCTGATGAACAAGACCCGCGCATCATGAGGAGGAGGC 267  
DB 1 GATGTGGAGCTGGGTCCTCTGAGTCTATGAACAACGAGAAAGATTATGGAACATGGG 60  
QY 268 GGGGCGCACTTCATCAACGCCCTTCGTGACCAACCCATGTGTGCTCCCTCAAGCTCCCTCC 327  
DB 61 GGGGCGCACTTCATCAATGCCCTTCGTGACCAACCCATGTGTGCTCCCTCAAGCTCCCTCC 120  
QY 328 ATCCCTCACTGGCAAGTACCTCCACACACACACACACACACACACACACACACACACACAC 387  
DB 121 ATGCTCACCGGGAAGTATGTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 180  
QY 388 TCGCCCTCTCGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 447  
DB 181 TCGCCCTCTCGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240  
QY 448 GGTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATACACAGGCTCTCTAGTGCCA 507  
DB 241 GGTACAGAACAGCCTTTTGGAAATACCTCAATGAATATAATGGCAGCTACATCCCC 300  
QY 508 CCGGCTGGAAGAGTGGCTCGGACTCTCTTAAAACTCCCGCTTTTATAACTTACACGCTG 567  
DB 301 CTTGGGTGGGAGATGGCTTGGATTAATCAAGATTTCTCGCTCTATTAATTAAGTCTT 360  
QY 568 TGTGGAACGGGGTGAAGAAAGACGGCTCCGACTTCAAGAGTATCTCAAGATTAAGTCAAGAC 627  
DB 361 TGTGCAATGGCATCAAGAAAGATGATGATTTGATGATGATGATGATGATGATGATGATGAT 420  
QY 628 CTCATCAACCAATGACAGCTGAGCTTCTCCGAGTCCCAAGATCTACCGGAGCAGG 687  
DB 421 TTAATCACTTAACGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480  
QY 688 CCAAGTCTCATGATCAGCATGAGCCGCCCCAGCCGCTTCAAGGATTCAGCCGCCA 747  
DB 481 CCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 748 TATTCAGGCTCTTCCCAAGCATCTCAGACATCAGCCGAGTCAACTAGCCGCC 807  
DB 541 TTTTCAAACTGTAACCCCAATGCTTCCCAACATTAATCTCTAGTTTAACTATGACCA 600  
QY 808 AACCCGGACAAACACTGTGATCATGCGCTACAGGGGCCCCATGAGCCCATCCACATGGA 867  
DB 601 AATATGATTAACACTGGATTATGCAATACAGGACCAATGCTGCCCATCCACATGGA 660  
QY 868 TTAACCAATGCTCCAGGGAAGGCTTGACAGCCCTCATGTCTGGTGGAGCATCCCATG 927  
DB 661 TTTTACAAACATTTACAGCGCAAAAGGCTTCCAGACTTGTGATGATGATGATGATGATGATGAT 720  
QY 928 GAGACGATTTTAAACATGCTGGTTGAGACGGCGAGCTGGACACACGCTACATCGTATAC 987  
DB 721 GAGAGCTGTATAACATGCTCTGAGACGGGGAGCTGGAGATTAATTAATTAATTAATTAAT 780  
QY 988 ACCGCGACACAGGTTTACATCGGCCAGTTTGGCTTGGTGAAGGGAATTCATGCCA 1047  
DB 781 ACCGCGACCATGTTTACATATGAGGAGTTTGGAGCTGTTGAGGAGGAAATTCATGCCA 840  
QY 1048 TATGAGTTTGAATCAGGTTCCGTTCTACGTGAGGGGCCCCAAAGTGAAGCGGCTGT 1107  
DB 841 TATGATTTTGAATATGCTGTGCTTTTATTTTATTTGTTGTTTGAAGTGAAGCAGGATCA 900  
QY 1108 CTGAATCCCAATCTCTCAATGATGACCTTGGCCCCACCATCTCTGAGCATTTGAGGC 1167  
DB 901 ATAGTCCCAAGATGCTTCTCAATGATGATTTGGCCCCCAGCATCTCTGATTTGCTGG 960  
QY 1168 CTGACATACCTCGGATATGAGCGGAATCTCACTCTCAAGCTCTGAGCAGGAGCGG 1227

DB 961 CTCGACACACCTCTCTGATGTGGACGGCAAGTCTGTCTCAAACTTCTGGACCCAGAAAG 1020  
QY 1228 CCGGTGAATCGGTTTCACTTTGAAAGAAAGATGAGGGTCTGGGGGAGCTCTTCTTGTG 1287  
DB 1021 CCAAGTAAACAGGTTTCCAAACAAACAGAAAGGCCAAATTTGGCGTGATACATTTCTCTAGTG 1080  
QY 1288 GAGAGAGGCAAGCTGTCTACCAAGAGAGACAATGACAAGGTGAGCCGCCAGAGAGAGAAC 1347  
DB 1081 GAAAGAGGCNAATTTCTAGTAAAGAGGAAGATCCAGCAAGATATCCACAGTCAAT 1140  
QY 1348 TTTCTGCCCAAGTACCAAGGTGTGAAGGACCTGTGTGTCAGGTGTGAGTACCAAGACGGCG 1407  
DB 1141 CACTTGCCTCCAAATATCAAGCGGTCAAGAACTATGCCAGCAGGCCAGGTACCAAGACAGCC 1200  
QY 1408 TGTGACGAGCTGGGACAGAAAGTGGAGTGTGGAGGACGCCACGCGGGAGCTGAAGCTG 1467  
DB 1201 TGTGAACACCGGGGCGAAGTGGCAATGCAATGAGATACATCTGGCAAGTTCGAAT 1260  
QY 1468 CATAGTGAAGGCGCCCATGCGGTGGGC-----GGCAGCAGAGCCCTCTCC 1515  
DB 1261 CACAAGTGAAGGACCCAGTGACCTGTCTACAGTCCGGCAGAGACGCGGAACCTCTAC 1320  
QY 1516 AACCTGTCGCCCAAGTACTACGGGAGG-----GCAGCGGCGCTGCACCTGTGAC 1566  
DB 1321 GCTCGCGCTTCCATGACAAAGAGTCAACGGCAATTTCTTGAGAAACAGGGGACTCCAAAGTAC 1440  
QY 1567 AGCGGGGACTTACAAGCTCAGCCTGGCGGACCGCGGAAAAAACTCTTCAAGAAAGATAC 1626  
DB 1381 AGCAGAAGCCAAAGAAAGAGTCAACGGCAATTTCTTGAGAAACAGGGGACTCCAAAGTAC 1440  
QY 1627 AAGCCAGCTATGTCCGAGTGTCTCTCACTCGCTCAGTGGCCATCGAGGTGAGCGGAGG 1686  
DB 1441 AAGCCAGATTTTCCATATCTCGGACAGACCTTCTTCTCGCTCGAATTTGAAGGTGAA 1500  
QY 1687 GTGTACACAGTGGCTGGGTGATGCCG-----CCAGCCCGGAAAACTCT 1731  
DB 1501 ATATATGACATAAATCTGGAAGAGAGAAATTTGCAAGTGTTCACCAAGAAACATT 1560  
QY 1732 ACCAAGCGGACCTGGCCAGGGGCCCCGTAGAGACCAAGATGACAAAGATGGTGGGAGCTTC 1791  
DB 1561 GCTAAGCGTCAATGATGAAGGCCCAAGGGGCCCAAGAGATCTCCAGGCTTCCAGTGTGCG 1620  
QY 1792 AGTGGCACTGGAGGCTTCCCGACTTCTCAGCGGC-----AACCCCAATAA 1839  
DB 1621 AACGGGCGAGGATGCTGGCAGATAGCAGCAAGCGCTGGGGGCCCACTTACACTGTCCGA 1680  
QY 1840 GTGACACATCGGTGTACATCTCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTAC 1899  
DB 1681 GTGACACACAAAGTGTATTTCTTCCCAATGACTCTATCCATTTGTGAGAGAGAACTGTAC 1740  
QY 1900 AAGTCCCTGCGCCCTGGAAGACCAACAGCTGCACTCGACCAAGAGATTGAAACCCCTG 1959  
DB 1741 CAATCGGCGAGGCTGGAGGAGCCATAAGGAGGATGATGACAAAGAGATTGAAGCTCTG 1800  
QY 1960 CAGAACAAATTAAGAACCTGGAGGAGTCCGAGGTCACTTGAAGAAAAAGCGGCGCAGAA 2019  
DB 1801 CAAGATAAAATTAAGAAATTTAAGAGAGTGAAGGACATCTGAAGAGAAAGAAAGCTGAG 1860  
QY 2020 GAATGTGACTGTCAAAATCAGTATCAGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2079  
DB 1861 GAATGTGACTGTCAAGTAAACAAAGCTTATTAATAAAGAGAAAGGAGTGAAGAAAGCAAGAG 1920  
QY 2080 GGTCTCAG-----TCTGCACTCTTTTTCAGGAAG-----GGCTTGAAGAGAGGAGCAAGGTG 2130  
DB 1921 AAATTAAGAGCCATCTTCAACCAATTCAGAGGAGGCTGCTCAGGAAGTAGATAGCAAACTG 1980  
QY 2131 TGGCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2190  
DB 1981 CAACTTTTCAAGGAGAACCAACCGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040  
QY 2191 AACAAAC---GACAGCTGAGCAGTCCAGGCTTCAAGTCTTCAACCCAGCAACCAAGCAGAC 2247  
DB 2041 AGGAAGGGGAGAGGTGACGCTTGGCTTCTTGGCTTCTTGGCTTCTTGGCTTCTTGGCTTCTTGGCT 2100

Qy	2248	TGSCAGACGGCGCTTTCTGAGCACTGGGGCTTTCTGTGCTCTGCAACGAGCGCAACAAT	2307
Db	2101	TGSCAGACAGCCCGTTCTGGAACTGGGATCTTTCTGTGCTTGCACGAGTTCTTAACAAT	2160
Qy	2308	AACACGCTACTGGTGCATGAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTT	2367
Db	2161	AACACCTACTGGTGTTCGTACAGTTAATGAGAGCGATAATTTCTTTCTGTGAGTTT	2220
Qy	2368	GCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCTACAGCTGATGAAATGCA	2427
Db	2221	GCTACTGGCTTTTGGAGTATTTTGATATGAATACAGATCCTTATCAGCTCACAAAATACA	2280
Qy	2428	GTGAACACACTGAGCAGGGATGTCCTCAACCGCTACACGTACAGTCTATGGAGCTGAGG	2487
Db	2281	GTSCACACGTTAGACAGAGGCATTTTGAATCAGTACAGTACACGTACAAGTAATGGAGCTCAGA	2340
Qy	2488	AGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACCTGMAACATGGACCTGGGACCTAAA	2547
Db	2341	AGCTGTCAAGGATATAAGCAGTGCACCCCAAGACCTTAAGATCTTGATGTTGGAAATAAA	2400
Qy	2548	GATGGAGGAAGCTATGAGCAATACAG	2573
Db	2401	GATGGAGGAAGCTATGACCTTACACAG	2426

## RESULT 7

RESOLUT. /  
US-09-668-673B-13  
/ Sequence 13, Application US/09668673B  
/ Patent No. 6562956  
/ GENERAL INFORMATION:  
/ APPLICANT: Emerson, Charles P  
/ APPLICANT: Dhoot, Gurtej K  
/ TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
/ TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
/ TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
/ FILE REFERENCE: PENN-0733  
/ CURRENT APPLICATION NUMBER: US/09/668,673B  
/ CURRENT FILING DATE: 2000-09-22  
/ PRIOR APPLICATION NUMBER: 60/155,738  
/ PRIOR FILING DATE: 1999-09-23  
/ NUMBER OF SEQ ID NOS: 22  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 13  
/ LENGTH: 1367  
/ TYPE: DNA  
/ ORGANISM: Mus sp.  
US-09-668-673B-13

Query Match	29.8%;	Score	1001.6;	DB	4;	Length	1367;
Best Local Similarity	86.2%;	Pred. No.	1.5e-247;				
Matches 1181;	Conservative	3;	Mismatches	162;	Indels	24;	Gaps
							6;
QY	1335	CCAGGAGAGAACTTTCTGCCCAAGTACACAGCTGTGAAGGACCTGTGTCAAGCTGTGA	1394				
Db	1	CCAGGAGGAGAACTTCTGTGCCCAAGTACACAGCGCGTGAAGGACCTGTGTCAAGCAGCTGA	60				
QY	1395	GTACACAGACGCGGTGTACAGCTGTGGACAGAAAGTGGCAGTGTGTGAGAGAGCCACGGG	1454				
Db	61	GTCACAGACAGCATCGGAAAGCTGTGGGCGAAGTGGCAGTGTGTGAGAGAGCGCTTCTCG	120				
QY	1455	GAAGCTGAAGCTGCATTAAGTGCAAAGGCGCCCATGCGGCT-----GGCGGCGACACAG	1505				
Db	121	CACGCTGAGCTGCACAAATGTAAAGGCCCATGCGGTTTGGTGGCGCGGTGGCACAG	180				
QY	1506	AGCCCTTCCAACTCTGTGTGCCCAAGTACTACGGGCGAGGCGACGAGGCGCTGCACCTGTGA	1565				
Db	181	AGCCCTTCTCAACCTGTGTGCCCAAGTATGACGGCCAGACGACGAGGCGCTGCACCTGTGA	240				
QY	1566	CA-----GCGGAGACTACAACTCAGCTCAGCTGCCCGGACGCCGGAATAAACTTTCAA	1616				
Db	241	CAGTGGCGGTGGAGGGGAGCTACAACTGGGCGCTGGGTGGACGCCG---TAAGCTCTTTAA	297				

Qy	1617	GAAGAAAGTCAAGGGCAGCTATGTTCGACGTGCTCTCATTCGGTCTCAGTGGCCATCGAGGT	1676
Db	298	GA AAAAGTATTAAGACAGAGCTATGCCCGAACCGCTCCATCCGTTCCGTGGCCATCGAGGT	357
Qy	1677	GGACGGCAGGGTGTACACGATAGCGCTGGGTGATGCGCCACGCCGAAACCTCACCAA	1736
Db	358	GGACGGTGAATATACCATAGAGGCTTGGATATCTGTGCTCAGCCCGCAACCTTAGCAA	417
Qy	1737	CGGGCATGGCCAGGGGCCCCCTGAGGACCAAGATGACAAAGATGTTGGGACCTTCAGTGG	1796
Db	418	CGCCGCACTGSGYACGGGGCCCTGAAGACCAAGATGACAAAGATGTTGGGACCTTCAGTGG	477
Qy	1797	CAC TGGAGGCTTCCCGAC TACTCAGCGCCACCCATTAAGATGACACATCGGTGCTA	1856
Db	478	TACTGGTGGCTTCCAGATTA TTTCTGCCCCCAATCCATCAAAGTAGCCCATCGGTGCTA	537
Qy	1857	CATCTTAGAAGACGACACAGCTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGCGAGGCTTG	1916
Db	538	CATCCTTGAAATGACACAGTCCAGTGGACCTTGGACCTGTATCAAGTCCCTGCGAGGCTTG	597
Qy	1917	GAAGACCAACAGCTCGACATCGACCACAGATTTGAAACCTTCGACGAA CAAAATTAAAGAA	1976
Db	598	GAAGACCAACAGCTCGACATCGACCATGAGATTCGAAACCTTCGACGAA CAAAATTAAAGAA	657
Qy	1977	CTTGAGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCGCCGACGAAGATGTGACTGTCACAA	2036
Db	658	CTTTCGAGAAGTCAGGGTCACTCTGAAGAAAGCGACCGGAAGATGTGACTGTCCATAA	717
Qy	2037	AATCAGCTACCAACCCAGCACAAGAGCCGCTCAAGCACAAGAGGTCCAGTCTGCATCC	2096
Db	718	AATCAGTTTACACAGCCACAACAAGGCCGCTCTCAAGCACAAGAGGTCCAGCCTGCACCC	777
Qy	2097	TTTCAGAAAGGGCCTGCAAGAGAAGCAAGGTGTGGCTGTTTGGGAGCAGAAAGCGCAA	2156
Db	778	TTTCAGAAAGGTCTGCAGAGAGAGCAAGGTGTGGCTGTCTGCGGGA -CAGAAACGCGAA	836
Qy	2157	GAAGAACTTCGCAAGCTGCTCAAAGCCTCGAGAACCAACGACAGTGCAGCATGCCAGG	2216
Db	837	GAAGAACTTCGCGCAA -CTGCTCAAAACGGCTGCAGAACACGATACGTGCAGCATGCC -GG	894
Qy	2217	CCTCAGCTGCTTACCCAGCAACAACAGCAGTGGCAGACGGCGCCTTTCTGACACATGGG	2276
Db	895	CCTCAGCTGCTTTACCCAGCAAAACCACTGGCAGACGGCGCCACTCTGCAGCGCTGGG	954
Qy	2277	GCCTTTCTGTGCTGCACAGCGCCAAACAAATAACACGTACTGCTGCATGAGGACCATCAA	2336
Db	955	GCCTTTCTGGCTGCACACAGCGCCAAACAACACGTACTGCTTGGAGGACCATAA	1014
Qy	2337	TGAGACTCAAAATTTCTCTCTGTGAATTTGGAACTGGCTTCTCTAGAGTACTTTGATCT	2396
Db	1015	TGAGACCCCAACTTCTCTTCTGGGAAATTTGCAACCGGCTTTCATAGAATACTTTGACCT	1074
Qy	2397	CAACACAGACCCCTTACACGCTCATGAATGACGTGAACACACCTGGACAGGGATGTCCTCAA	2456
Db	1075	CAGTACAGACCCCTTACACGCTCATGAATGACGTGAACACACCTGGACAGGGATGTCCTCAA	1134
Qy	2457	CGAGCTACAGTACAGCTCATGAGACTGAGGAGCTGCAAGGGTTTACAAGCAGTGTAAACC	2516
Db	1135	CCAACTGCACGTGACGCTCATGAGACTAAGGAGCTGTAAAGGCTACAAGCAGTGCACCC	1194
Qy	2517	CGGGACTCGAAACATGGACCTTGGGACTTAAAGATGGAGGAGCTTATGACCAATACAGGCA	2576
Db	1195	CGGGAACCCGCAACTTGGACCTTGGGCTTTAGAGACGGAGGAGCTTATGACCAATACAGGCA	1254
Qy	2577	GTTTTCAGCGCTGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGCAACT	2636
Db	1255	GTTTTCAGCGCTGAAATGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGCAAGCT	1314
Qy	2637	GTGGGAAGCTGGGAAGGTTTAAAGAAACAACAGAGGTGGACCTTCCAAAAC	2686
Db	1315	ATGGGAAGGTTGGGAAGGTTAGCGGCTATGAGAGAGGAACTTCCAAAAC	1364

## RESULT 8

US-09-810-347-1  
; Sequence 1, Application US/09810347  
; Patent No. 6461847  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: CL001169  
; CURRENT APPLICATION NUMBER: US/09/810,347  
; CURRENT FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1799  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-347-1

Query Match 17.2%; Score 578.2; DB 4; Length 1799;  
Best Local Similarity 73.1%; Pred. No. 1.1e-138;  
Matches 742; Conservative 0; Mismatches 273; Indels 0; Gaps 0;  
QY 98 TGTGTGGTGGAGCTGGCTTCTCTGTCGACACCAACCGCTTGAAGCAGGTTTCAGAGG 157  
DB 156 TGTGTGGTGGAGCTGGCTTCTCTGTCGACACCAACCGCTTGAAGCAGGTTTCAGAGG 215  
QY 158 ACCGAGGAACATCCGCCCAACATCATCTCTGTGTGTGACGACGACCAAGGATGGAGC 217  
DB 216 AACGAAAAAATCCGACCCCAACATATTCTGTGTGTGACGATGATCAAGATGGAGC 275  
QY 218 TGGTTCATGCGAGTGATGAACAGACCGCGGCGATCATGAGAGCGGCGGCGCACT 277  
DB 276 TGGGTGCTTCCGCAAGTCATGAACAAACGAGAAAGATTATGGAACATGGGGGGGCCACCT 335  
QY 278 TCATCAACGCTTCTGTGACACACACCATGTGTGCTTCCCTCAGCTCTCTCCATCTCACTG 337  
DB 336 TCATCAATGCTTGTGACTACACCATGTGTGCTTCCCTCAGCTCTCTCCATCTCACTG 395  
QY 338 GCAAGTAGCTCCCAACACCAACACCTTACCAACCAATGAGAACTGCTCTCTGCGCTCT 397  
DB 396 GGAAGTAGTGCAACATCAAAATGTCTACACCAACCAAGAACTGCTCTCTGCGCTCT 455  
QY 398 GGCAGGACACGACGAGCGGCGACCTTTCGCGTGCTACATAGCACTGCTGCTACCGA 457  
DB 456 GGCAGGCGATGATGAGCCTCGGACTTTTGTCTGTATCTTTACCAACATGCGCTACAGAA 515  
QY 458 CAGCTTTCTTTCGGGAAGTATCTTTAATGAATACAAACCGCTCTACGTCGCCACCGGCTGA 517  
DB 516 CAGCTTTTTCGGAATACTCTCAATGAATATTAATGGCAGCTACATCCCTCTGGGTGGC 575  
QY 518 AGAGTGGTTCGAGCTCTTAAACCTCCGCTTTTAACTACAGCTGTGTGCGAAGC 577  
DB 576 GAGATGGCTTGGATTAAATCAAGAACTTCTGCTTCTTAAATTAATTAATTAATTAAT 635  
QY 578 GGTGAAAGAAACACGCGCTCCGACTTCTTCAAGGATTACCTTCAAGACCTCATCACCA 637  
DB 636 GCATCAAGAAAGCATGGATTGATTAATGCAAGGACTTCTACAGACTTAATCACTA 695  
QY 638 ATGACAGCGTGTGCTTTCGCGCAGCTCCAGAGATGTACCGGACACAGGCGAGCTCTCA 697  
DB 696 ACAGAGCATTAATTAATCTTCAAAATGTCTAAGAGAAATGTATCCCATAGGCGCGTATGA 755  
QY 698 TGTGTCTACGCCATGACGCGCCCTTGGAGTTTCAAGGATTCAGCCCAACATATTTCAGGCG 757  
DB 756 TGTGTATCAGCCAGCTGTGCGCCCAACGCGCCCGAGGACTACGCCCCACAGTTTCTAATC 815  
QY 758 TCTTCCCAACGCACTCAGCAATCAACGCGGAGCTTCAACTACGCGCCCAACCGGACA 817  
DB 816 TGTACCCCAATGCTTCCCAACACATAACTCTAGTTTATAACTATGATGACCAATATGGATA 875  
QY 818 AACACTGGATCATGGGCTACACGGGGCCCATGAGGCCCATCCACATGGGAATTCACCAACA 877

DB 876 AACACTGGATTATGCACTACACAGCAATGCTGCCCATCCACATGGAATTTACCAACA 935  
QY 878 TGTCTCAGCGAAGCGCTTGCAGACCTCATGTGCGTGGAGCGACTCCATGGAGACGATTT 937  
DB 936 TTTTACAGCGCAAAAGGCTCCAGACTTTTGATGTGTCAGTGGATGATTTCTGTGGAGAGGCTGT 995  
QY 938 ACAACATGCTGTGTTGACACGGCGAGCTGGACAACAGTACATCGTATACACCGCGGACC 997  
DB 996 ATAACATGCTGTGGAGACCGGGGAGCTGGAGATCTTACATCATTTTACACCGCGGACC 1055  
QY 998 AGGTTTACCACATCGGCCAGTTTGGCTGTGTAAGAGGAAATCCATGCCATATGAGTTTG 1057  
DB 1056 ATGTTTACCATATTTGGCAGTTTGGACTGTGTAAGGGGAAATCCATGCCATATGACTTTG 1115  
QY 1058 ACATCAGGTCCTCGTCTTCTACGTGAGGGGCCCCACGTGGAGCGGCTGCTCA 1112  
DB 1116 ATATCGTGTGCTTTTATTCGTGTCCTCAAGTGTAGAACCCAGGATCAATGTA 1170

## RESULT 9

US-09-404-879A-315/c  
; Sequence 315, Application US/09404879A  
; Patent No. 6468546  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Algate, Paul A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462C2  
; CURRENT APPLICATION NUMBER: US/09/404,879A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 393  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 315  
; LENGTH: 441  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-404-879A-315

Query Match 13.1%; Score 441; DB 4; Length 441;  
Best Local Similarity 100.0%; Pred. No. 9.6e-104;  
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2995 CAAAAATGCAATTTTTCGTATCAAAAAGTCAACCTTCCCTCCCGAGAGCTCACAA 2954  
DB 441 CAAAAATGCAATTTTTCGTATCAAAAAGTCAACCTTCCCTCCCGAGAGCTCACAA 382  
QY 2955 AGGAAACGAGAGAGAGAGAGAGAGATTTCTTTGGAAATTTCTCCCAAGGCGAAG 3014  
DB 381 AGGAAACGAGAGAGAGAGAGAGATTTCTTTGGAAATTTCTCCCAAGGCGAAG 322  
QY 3015 TCATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTTAAATCTCTTATTTCT 3074  
DB 321 TCATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTTAAATCTCTTATTTCT 262  
QY 3075 TGGTTCTCAAGAGAGGAACTAAGACGAGCAGACAGGACGAGAGAGCTGGAAG 3134  
DB 261 TGGTTCTCAAGAGAGGAACTAAGACGAGCAGACAGGACGAGAGAGCTGGAAG 202  
QY 3135 CAGTCAGAGACGTTTGGACAATGAGTCAAGTACAGTACCAAAAAGAGATGACATTTACCTAGCAC 3194  
DB 201 CAGTCAGAGACGTTTGGACAATGAGTCAAGTACAGTACCAAAAAGAGATGACATTTACCTAGCAC 142  
QY 3195 TATAAACCTTGGTTCCTCTGAAGAACTCCCTTCAATGTTATATATGACTATTTTACAT 3254  
DB 141 TATAAACCTTGGTTCCTCTGAAGAACTCCCTTCAATGTTATATATGACTATTTTACAT 82  
QY 3255 GTATCAACATGGGAATTTTAGGGGAACCTAATAAGAAATCCCAATTTTTCAGGAGTGT 3314  
DB 81 GTATCAACATGGGAATTTTAGGGGAACCTAATAAGAAATCCCAATTTTTCAGGAGTGT 22

QY 3315 GGTGTCAATAAACGCTCTGTG 3335  
Db 21 GGTGTCAATAAACGCTCTGTG 1

## RESULT 10

US-09-668-673B-10  
; Sequence 10, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 494  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-09-668-673B-10

Query Match 11.3%; Score 381.6; DB 4; Length 494;  
Best Local Similarity 89.5%; Pred. No. 1.9e-88;  
Matches 444; Conservative 0; Mismatches 49; Indels 3; Gaps 3;  
QY 2003 AGAAAGCGGCCAGAGNATGTGACTGTCAAAAATCAGTACACACCAGCAAAAG 2062  
Db 1 AGAAGAGCGCCGGAAGATGTGCTGCCATAAAATCAGTTACACAGCCCAACACAAAG 60  
QY 2063 GCGGCTTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGG 2122  
Db 61 GCGGCTTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGG 120  
QY 2123 ACAAGTGTGGTGTGGGAGCAGAGAGCGCAAGAACTCGCAAGTGTCTCAAGC 2182  
Db 121 ACAAGTGTGGTGTGGGAGCAGAGAGCGCAAGAACTCGCAAGTGTCTCAAGC 178  
QY 2183 GCGTCAGAACACACAGCTGACATGCGCCAGGCTCACTGCTTACCCACAGCAACC 2242  
Db 179 GCGTCAGAACACACAGCTGACATGCGCCAGGCTCACTGCTTACCCACAGCAACC 237  
QY 2243 AGCACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGGCGCTTGACACAGCGCA 2302  
Db 238 ACCACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGGCGCTTGACACAGCGCA 297  
QY 2303 ACAATAACACGTACTGTGTGATGAGGACATCATGAGCTCAAAATTCCTTCTGTG 2362  
Db 298 ACAATAACACGTACTGTGTGATGAGGACATCATGAGCTCAAAATTCCTTCTGTG 357  
QY 2363 AATTGCAACTGGCTTCTAGAGTACTTGTATCTCAACACAGACCCCTACACAGCTGATGA 2422  
Db 358 AATTGCAACTGGCTTCTAGAGTACTTGTATCTCAACACAGACCCCTACACAGCTGATGA 417  
QY 2423 ATGCAGTGAACACACTGACAGAGGATGCTTCAACAGCTACAGTACAGTCTATGAGC 2482  
Db 418 AGCGGTGAACACACTGACAGAGGATGCTTCAACAGCTACAGTACAGTCTATGAGC 477  
QY 2483 TGAGGAGTGAAGG 2498  
Db 478 TAAGGAGTGAAGG 493

## RESULT 11

US-09-668-673B-9  
; Sequence 9, Application US/09668673B  
; Patent No. 6562956

; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 466  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-09-668-673B-9

Query Match 11.2%; Score 375.8; DB 4; Length 466;  
Best Local Similarity 90.6%; Pred. No. 5.6e-87;  
Matches 423; Conservative 0; Mismatches 42; Indels 2; Gaps 2;  
QY 1885 GACCTGACCTGTACAGTCCCTGACGGCTGGAAGACCAACAGCTGCATCGACCCAC 1944  
Db 2 GACCTGACCTGTACAGTCCCTGACGGCTTGAAGACCAACAGCTGCATCGACCCAT 61  
QY 1945 GAGATTGAACACCTGCGAGAACAAATTAAGAACCTGAGGAAAGTCCGAGGTCACTGAAG 2004  
Db 62 GAGATTGAACACCTGCGAGAACAAATTAAGAACCTGAGGAAAGTCCGAGGTCACTGAAG 121  
QY 2005 AAAAGCGGCCAGAGATGTGACTGTCAAAAATCAGTACACACCAGCAAAAGC 2064  
Db 122 AAGAGCGACCGGAAGATGTGACTGCCATAAAATCAGTTACACAGCCCAACAAAGC 181  
QY 2065 CGCTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGGAC 2124  
Db 182 CGCTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGGAC 241  
QY 2125 AAGGTGTGCTGTGGGAGCAGAGCGCAAGAACTCCGCAAGTGTCTCAAGCGC 2184  
Db 242 AAGGTGTGCTGTGGGAGCAGAGCGCAAGAACTCCGCAAGTGTCTCAAGCGC 299  
QY 2185 CTGCAGAACACACAGCTGACGATGCCAGGCTCAAGGCTTACCCACAGCAACCG 2244  
Db 300 CTGCAGAACACACAGCTGACGATGCCAGGCTTACCCACAGCAACCG 359  
QY 2245 CACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGGCGCTTGACACAGCGCCAC 2304  
Db 360 CACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGGCGCTTGACACAGCGCCAC 419  
QY 2305 AATAACACGTACTGTGTGATGAGGACCATCAATGAGACTCAAAATTT 2351  
Db 420 AATAACACGTACTGTGTGATGAGGACCATCAATGAGACTCAAAATTT 466

## RESULT 12

US-09-484-970B-70  
; Sequence 70, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne G.  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 70  
; LENGTH: 3871  
; TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6426186 481723.1CB1

NAME/KEY: unsure

LOCATION: 2910-2932, 3791, 3794, 3813, 3815, 3827, 3843, 3850-3851, 3860

OTHER INFORMATION: a, t, c, g, or other

US-09-484-970B-70

Query Match 11.0%; Score 369; DB 4; Length 3871;

Best Local Similarity 58.6%; Pred. No. 9.1e-85;

Matches 804; Conservative 0; Mismatches 505; Indels 62; Gaps 7;

1265 TCTGGCGGACTCCTTCTTGTTGAGAGAGCAAGCTGTACACAAGAGAGACAATGACA 1324

6 TTTGGCGGTGATACATTCTTAGTGGAAAGAGGCAAAATTTACGTAAGAGGAAATCCCA 65

1325 AGTGGAGCGCCAGGAGGAGAACTTTCTGCCCAAGTACCAGCGTGTCAAGGACCTGTGTC 1384

66 GCAAGATATCCACAGTCAATCACTTGGCCCAATATGACGGGTCAAGAACTATGCC 125

1385 AGCTGTGTAGTACACAGCGCGTGTGAGCAGTGGGACGAAAGTGGCAGTGTGGAGG 1444

126 AGCAGGCGAGTACACAGACGCTGTGAACAACCGGGCGAAGTGGCAATGCATTGAGG 185

1445 ACGCCAGCGGGAAGCTGAAGTGCATAGTGCAGAGGCGCCNCTCGCGCTGGCC 1497

186 ATACATCTGGCAAGCTTCGAATTCACAAGTGTAAAGGACCCAGTGACCTGCTCACAGTCC 245

1498 -----GGCAGCAGAGCCCTCTCCAACTCGTGGCCCAAGTACTACGGGCGAGG----- 1543

246 GGCAGAGCACGCGAACTCTACGCTCGCGCTCCATGACAAAGACAAAGAGTGCAGTT 305

1544 GCAGCGAGGCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCTGGCGCGAGCGGGA 1603

306 GTAGGAGTCTGGTTACGTTGCCAGCGAAGGCCAAAGAAAGTCAACGGCAATTTTGA 365

1604 AAAAAGCTCTTCAAGAAAGTACAAGCGCCAGCTATGTCGCGAGTCCATCCGCTCAG 1663

366 GAAACCGGGGACTCCAAAGTACAGCCAGTTTGTCCATCTCGGCAGACAGTTCTCT 425

1664 TGGCCATCGAGTGGAGCGGCGGTGTACCAAGTGGCCCTGGGTGATGCGG----- 1714

426 TGCCCGTCGAATTTGAAGGTGAATATATGACATAAATCTTGGAAAGAAAGAAATTCG 485

1715 -----CCGACCCCGAACTCCACCAAGCGCACTGGCCAGGGGCGCTGAGACCAAG 1768

486 AAGTTTGCACACCAAGAAACATTCCTAAGCTCATGATGAAGGCCCAAGGGGCCAAGAG 545

1769 ATGACAAAGATGTGGGGACTTCAGTGGCACTGGAGGCGCTTCCGCACTACTCAGCGCGC- 1827

546 ATCTCCAGGCTTCCAGTGGTGGCAACAGGGGCGAGGATGCTGGCAGATAGCAGCAAGCGCG 605

1828 -----AACCCATTAAGTGCACATCGTGTCTACCTCTAGAGACGACAG 1876

606 TGGGCCCACTTACCTGTCGAGTGCACACAAGTGTATTCTTCCCAATGACTCTA 665

1877 TCCAGTGTGACTGGACCTGTACAAGTCCCTGCGAGGCGCTGGAAGACCAACAGCTGCACA 1936

666 TCCATTGTGAGAGAGACTGTACCAATCGCCAGAGCGGTGGAAGACCATAGGCATACA 725

1937 TCGACACAGAGATTGAACCTGTGACGAAACAAATTAAGAACCTGAGGGAAGTCCGAGGTC 1996

726 TTGACAAAGAGATTGAAGCTCTGCAAGATAAATTAAGAAATTTAAGAGAAAGTGAAGGAC 785

1997 ACCTGGAAGAAAGCGGCCAGAGAAATGTGACTGTCAAAATTCAGCTACCAACCCAGC 2056

786 ATCTGAGAGAGAGAGCTTGAAGATGTAGCTGCGAGTAAACAAGCTTTAACAATAAG 845

2057 ACAAGGCCCGCTCAAGC-----ACAGAGGCTCCAGTCTGCATCCTTTCAAGGAG 2106

846 AGAAGGTGTAAAGAAAGCAAGAGAAATTAAGAGGCACTCTTACCCATTTCAAGGAGGCTG 905

2107 GGCTTGCAAGAGAGCAAGGTGTGGCTGTGGCGGAGCAGAGCGCAAGAGAACTC 2166

906 CTCAGGAGTAGATAGCAAACTGCAACTTTTTCAGGAGAACACACCTGTAGAGGAGAG 965

2167 CCAGAGCTGTCAAGCGCTGTGAGAAACA---CGACAGTGCAGCATGCC-AGGCTTCC 2222

966 GAGAGGAGGAGAGAGAGCGGAGGAGGAGGAGAGTGCAGCTGCTGGGCGCTCAC 1025

2223 GTGCTTCCACACGACCAACACGACTGCGCAGACGCGCTTCTGGACACTGGGCGCTTT 2282

1026 TTGCTTCCAGGATGACAAACCACTGCGCAGACAGCCCGTTCTGGAACTTGGGATCTTT 1085

2283 CTGTCCTGCGACCGCCAGCCAAACAACAACGTAAGTGTGATGAGGACCAATCAATGAGAC 2342

1086 CTGTCCTGCGAGGTTCTTAACAATAACACCTACTGTGTGTTGGTACAGTAATGAGAC 1145

2343 TCACAAATTCCTCTCTCTGTAATTTGCACTGGCTTCTAGAGTACTTTGATCTCAACAC 2402

1146 GCTAATTTCTTTCTGTGAGTTTGTACTGGCTTTTGGAGTATTTGATATGAATAC 1205

2403 AGACCCCTACCAAGCTGATGAATGCAAGTGAACACACACTGGACAGGATGCTTCAACCACT 2462

1206 AGATCCTTATCAGCTCAACAATACAGTGCACACGCTAGACAGGCAATTTTGAATCAGCT 1265

2463 ACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTACAAGCAGTGTAAACCCCGGAC 2522

1266 ACAGTACAGCTCAATGAGGCTCGAAGCTGTCAAGGATATAAGCAGTGCACCCCAAGACC 1325

2523 TCGAAACATGCACTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAG 2573

1326 TAAGATCTTGTATTTGGAATAAAGATGGAGGAGTATGACCTACACAG 1376

# RESULT 13

US-09-668-673B-12

; Sequence 12, Application US/0968673B

; Patent No. 6562956

; GENERAL INFORMATION:

; APPLICANT: Emerson, Charles P

; APPLICANT: Dhoot, Gurtej K

; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF

; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR

; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS

; FILE REFERENCE: PENN-0733

; CURRENT APPLICATION NUMBER: US/09/668,673B

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: 60/155,738

; PRIOR FILING DATE: 1999-09-23

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 459

; TYPE: DNA

; ORGANISM: Mus sp.

US-09-668-673B-12

Query Match 10.9%; Score 368; DB 4; Length 459;

Best Local Similarity 87.9%; Pred. No. 5.6e-85;

Matches 401; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

2231 CCCACGACCAACAGCACTGGGACAGCGGCGCTTTCTGGACACTGGGCGCTTTCTGTGCT 2290

1 CCCACGACCAACCACTGGGACAGCGGCGCACTCTGGAGCTGGGCGCTTTCTGTGCT 60

2291 GCACGAGCCCAACAATAACAGTACTGTGTGATGAGGACCAATGAGCTCAAAATT 2350

61 GCACGAGCCCAACAACAACAGTACTGTGTGATGAGGACCAATGAGGACCAAACT 120

2351 TCCTCTTTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCT 2410

121 TCCTCTTTCTGGAATTTGCAACCGGCTTCAAGATACCTTTGACCTCAGTACAGACCCCT 180

2411 ACCAGCTGATGAATGAGTGAACACACTGAGGAGGATGCTCTCAACCAAGCTACAGTAC 2470



Db 181 ACAGCTGATGAACCGGGTGAACACACTGGACAGGACGCTCTTAACCAACTGCACGTGC 240  
QY 2471 AGCTCATGAGCTGAGGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCGGACTCCGAACA 2530  
Db 241 AGCTCATGAGCTAAGGAGCTGTAAAGGCTACAAGCAGTGTAAACCCGGACCCCAACA 300  
QY 2531 TGGACCTGGGACTTAAAGATGGAAGGAGTATGAGCAATACAGGAGTTTCAGCGTCGAA 2590  
Db 301 TGGACCTGGGCTTAGAGACGGAGGAGTATGAGCAATACAGGAGTTTCAGCGTCGAA 360  
QY 2591 AGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGNAGGCTGG 2650  
Db 361 ATGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGNAGGCTGG 2650  
QY 2651 AGGTTTAAAGAAACACAGAGGTTGGACCTCCAAAAAC 2686  
Db 421 AAGGCTAAGCGCCATAGAGAGAGAACTCCAAAAAC 456  
RESULT 14  
US-09-668-673B-8/c  
; Sequence 8, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 538  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-09-668-673B-8  
Query Match 10.3%; Score 347.4; DB 4; Length 538;  
Best Local Similarity 82.1%; Pred. No. 1.2e-79;  
Matches 444; Conservative 0; Mismatches 76; Indels 21; Gaps 3;  
QY 1335 CCAGGAGGAGAACTTTCGCCCAAGTACAGCGTGTGAAGACCTGTGTACAGCGTGTGA 1394  
Db 538 CCAGGAGGAGAACTTTCGCCCAAGTACAGCGGCTGAAGGACCTGTGTACAGCGTGTGA 479  
QY 1395 GTACCAGACGGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTGAGGAGCGGCGG 1454  
Db 478 GTACCAGACAGCATGCGAACAGCTGGGCGCAGAGTGGCAGTGTGTGAGGAGCGGCTTCTGG 419  
QY 1455 GAAGCTGAAGCTGCATTAAGTCAAGGGCCCATCGGCT-----GGCGCGCAGCAG 1505  
Db 418 GACGCTGAAGCTGCACAAATGTAAAGGGCCCATCGGCTTTGGTGGCGCGCTGGCAGCAG 359  
QY 1506 AGCCCTCTCCAACTCGGCCCAAGTACTAGCGGAGGCGGACGAGCGCTGCACCTGTGA 1565  
Db 358 AGCCCTCTCCAACTCGGTCGCCCAAGTATGACGGCGCAGACGAGCGCTGCACCTGTGA 299  
QY 1566 CA-----GCGGGGACTACAAGCTCAGCGCTGGCGGACCGCGGAAAAAACTCTTCAA 1616  
Db 298 CAGTGGCGGTGGAGGGGACTACAAGCTGGGCTGGCTGGAGCGCG-----TAAGCTCTTTAA 242  
QY 1617 GAAAGAGTACAGGCGCAGCTATTCGCGAGTTCGCTTCATTCGCTCAGTGGCCATTCGAGGT 1676  
Db 241 GAAAAAGTATAAGACCAAGTATGCCCGGAAACCGCTCCATCCGTTCCGTCGCGCATCGAGGT 182  
QY 1677 GGACGGCAGGCTGTACCAAGCTAGGCTGGGTGTGCGCGCCAGCGCCAACTCACCAG 1736

Db 181 GGACGCTGAGATATACCACGCTAGGCTTGGATACTGTGCTCAGCCCCGCAACCTTAGCAA 122  
QY 1737 GCGGCACTGCCAGGGGCCCCCTGAGACCAAGATGACAGGATGGTGGGACTTTCAGTGG 1796  
Db 121 GCGGCACTGCTCAGGGGCCCCCTGAGACCAAGATGACAGGATGGTGGGCAAGTTTTCAGTGG 62  
QY 1797 CACTGAGGCGCTTCCCGACTACTCAGCGGCAACCCCATTTAAAGTGCACATCGGTGCTA 1856  
Db 61 TACTGTGGCTTCCAGATTATTCTGCCCCCAATCCATCAAAAGTACCCTCAGGTGCTA 2  
QY 1857 C 1857  
Db 1 C 1  
RESULT 15  
US-09-668-673B-11  
; Sequence 11, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 436  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-09-668-673B-11  
Query Match 8.8%; Score 297.4; DB 4; Length 436;  
Best Local Similarity 83.4%; Pred. No. 7.6e-67;  
Matches 366; Conservative 0; Mismatches 61; Indels 12; Gaps 2;  
QY 1501 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTAGGGCGAGGCGAGGCGCTGCACC 1560  
Db 1 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTAGGGCGAGGCGAGGCGCTGCACC 60  
QY 1561 TGTGACA-----GCGGGGACTACAAGCTCAGCGCTGGCGGACCGCGGAAAAAACTC 1611  
Db 61 TGTGACAGTGGCGGTGGAGGGGACTACAAGCTGGGCGCTGGCTGGAGCGCG---TAAGCTC 117  
QY 1612 TTCAGAAGAAGTACAGCGCAGCTATGTCCGAGTCCGCTCCATCCGCTCAGTGGCGCATC 1671  
Db 118 TTTAAGAAAAAGTATAAGACCAAGCTATGCCGGAACCGCTCCATCCGTTCCGTCGCGCATC 177  
QY 1672 GAGGTGGACGGCAGCGGTGTACCACTAGCGCTGGGTGTATGCCGCCAGCCCGCAACCTC 1731  
Db 178 GAGGTGGACGGTGCAGATATACCACTAGCTTGGATCTGTGCTCAGCCCGCGCAACCTT 237  
QY 1732 ACCAAGCGGCACTGGCCAGGGGCCCCCTGAGGACCAAGATGACAAAGATGGTGGGCACTTC 1791  
Db 238 AGCAAGCGGCACTGGCCAGGGGCCCCGTGAAGACCAAGATGACAAAGATGGTGGGCACTTC 297  
QY 1792 AGTGGCACTGGAGGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGCACATCGG 1851  
Db 298 AGTGGTACTGGTGGCTTCCAGATTATTCTGCCCCCAATCCATCAAAAGTGAACCATCGG 357  
QY 1852 TGCTACATCTTAGAGAACGACACAGTGTGACCTTGGACCTGTACAAGTCCCTGCAG 1911  
Db 358 TGCTACATCTTAGAGATGACACAGTCCAGTGGCACTTGGACCTGTACAAGTCCCTGCAG 417  
QY 1912 GCTTGAAGACCAAGC 1930  
Db 418 GCTTGAAGACCAAGC 436

Search completed: February 14, 2004, 19:42:07  
Job time : 235 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 13:53:07 ; Search time 6194 Seconds  
(without alignments)  
13195.989 Million cell updates/sec

Title: US-10-025-966A-4

Perfect score: 3363

Sequence: 1 gactccgcgcatcccaaga.....aaagaaaaaa 3363

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estinu.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estrio.\*

8: em\_estci.\*

9: gb\_estli.\*

10: gb\_est2.\*

11: gb\_est3.\*

12: gb\_est4.\*

13: gb\_est5.\*

14: gb\_estfun.\*

15: em\_estom.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2093.6	62.3	3704	11 AK036685	AK036685 Mus muscu
2	2093.6	62.3	3759	11 AK034712	AK034712 Mus muscu
3	2092	62.2	3687	11 AK028874	AK028874 Mus muscu
4	2092	62.2	3734	11 AK081643	AK081643 Mus muscu

5	1430.8	42.5	2685	11 AK049170	AK049170 Mus muscu
6	1062.4	31.6	4564	11 AK045002	AK045002 Mus muscu
7	1062.4	31.6	4513	11 AK078756	AK078756 Mus muscu
8	1060.8	31.5	4428	11 AK028285	AK028285 Mus muscu
9	856.4	25.5	1803	11 AK008108	AK008108 Mus muscu
10	838.2	24.9	1201	9 AL537332	AL537332 AL537332
C	829.4	24.7	1201	9 AL574439	AL574439 AL574439
11	829.4	24.2	948	13 BU956086	BU956086 AGENCOURT
12	812.8	23.7	1066	13 BU956086	BU956086 AGENCOURT
13	799	23.7	885	13 BU956086	BU956086 AGENCOURT
14	798.6	23.7	885	13 BU956086	BU956086 AGENCOURT
15	791.4	23.5	904	13 BU956086	BU956086 AGENCOURT
16	788.4	23.4	861	14 CD106296	CD106296 AGENCOURT
17	786.4	23.4	933	13 BX370010	BX370010 BX370010
18	777.6	23.1	916	12 BI911672	BI911672 603064857
19	760.8	22.6	899	14 CA489747	CA489747 AGENCOURT
20	754.8	22.4	843	13 BU858200	BU858200 AGENCOURT
21	749.2	22.3	874	13 BU858200	BU858200 AGENCOURT
22	748.6	22.3	2703	11 AK037389	AK037389 Mus muscu
23	742.8	22.1	899	13 BU858200	BU858200 AGENCOURT
24	725	21.6	856	13 BU858200	BU858200 AGENCOURT
25	722.8	21.5	967	10 BG386102	BG386102 602455276
C	721.4	21.5	890	13 BX411283	BX411283 BX411283
C	719	21.4	730	14 CB850913	CB850913 UI-CF-EN1
27	713.2	21.2	941	12 BG910212	BG910212 602805669
28	709.2	21.1	826	12 BU908810	BU908810 603066272
29	709.2	21.1	916	13 BU171062	BU171062 AGENCOURT
30	704.4	20.9	982	13 BU930803	BU930803 AGENCOURT
31	700.6	20.8	982	13 BU930803	BU930803 AGENCOURT
C	697.8	20.7	736	13 BU48044	BU48044 UI-H-EU1-
C	697	20.7	728	12 BM985096	BM985096 UI-CF-RC1
34	696.2	20.7	739	13 BU48044	BU48044 UI-H-EU1-
35	695.8	20.7	981	13 BU48044	BU48044 UI-H-EU1-
36	694.2	20.6	880	13 BU927919	BU927919 AGENCOURT
C	693.6	20.6	1201	9 AL566767	AL566767 AL566767
C	693	20.6	705	13 BU46043	BU46043 UI-H-EU1-
C	692.8	20.6	1201	13 BU461722	BU461722 BX461722
C	692.6	20.6	1045	10 BG036300	BG036300 602327081
C	689.4	20.5	703	13 BU181056	BU181056 UI-H-EU0-
C	688.8	20.5	703	12 BM981713	BM981713 UI-CF-EN1
43	685	20.4	753	10 BF530592	BF530592 602071865
44	682.8	20.3	755	10 BG248293	BG248293 602400329
45	676.8	20.1	704	12 BM930827	BM930827 UI-E-EU1-

## ALIGNMENTS

RESULT 1

AK036685

LOCUS

DEFINITION

AK036685 3704 bp mRNA linear HTC 05-DEC-2002  
Mus musculus adult male bone cDNA, RIKEN full-length enriched  
library, clone:9830162M14 product:weakly similar to  
N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert  
sequence.

ACCESSION

AK036685

VERSION

AK036685.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1

Carminci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

MEDLINE

11042159	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
3	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
20530913	Genome Res. 10 (11), 1757-1771 (2000)
11076861	20530913
4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Hill, D., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Lee, N. H., Lyons, P., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Mombaerts, P., Nordons, P., Marchionni, L., Mashima, J., Mazza, R., Sakamoto, N., Sasaki, H., Ring, B., Schönbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Sato, K., Schönbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
5	Functional annotation of a full-length mouse cDNA collection
21085660	Nature 409 (6921), 685-690 (2001)
11217851	21085660
5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
6	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
21085660	Nature 420, 563-573 (2002)
11217851	21085660
5	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
6	Direct Submission
21085660	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
7	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
8	Please visit our web site for further details.
9	URL: http://genome.gsc.riken.go.jp/
10	URL: http://fantom.gsc.riken.go.jp/

FEATURES	Location/Qualifiers
source	1. .3704
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:9830162M14"
	/db_xref="taxon:10090"
	/clone="9830162M14"
	/sex="male"
	/tissue type="bone"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
CDS	210. .2837
	/note="unnamed protein product; putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPTR Q90XB6, evidence: FASTY, 64.2% ID, 99.5% length, match=2601)"
	/codon_start=1
	/protein_id="BAC29534.1"
	/db_xref="GI:26331608"
	/translation="WAPGLPLWLLSTALLSLLAGSAPLSPRLKGRQRRNRNP NIILVLDQDVELGSMQVMMKTRIMEQGGAFINAVFTPMCCPSSSILTGKIVH NNTYTNENSCSPSQACHSRTPAVLNSTGYRTAFKGLNYNEYSYVPGKWE VGLKNSRPNYTLCRNGVKEHSDYSTYLDLTINDSVSFRTSKMYPHRPVLM VLSHAHPGEPDAPQYSLPNAHQHTPSVYAPNDKWMYTGPKYIHKMET NMLQKRLQTLMSVDDSEIYDMLVEGELDNTILYTDHGYHIGQGLYKGSMP YEDIRPFYVRGPNVAGSLNPHIVINDLAPTLIDTAGDIADMDGSKSLKIDS EEPVNRFLHKKLVNRDLSFLVERGLKLRREGDKVNAENFLPKYQKVDLCQRAE YTAECQLQWQVEDASGTLKHKCKPGRFGGSGSRALSNLVPKVDGSSSEAC CSQGGDYKGLAGRLKFKKTKYTSVARNRSIRVAIEVDGEIYHVLDTVPQPN LSKHPGAPDQDDKDGSPGCTGLPDYAPNPVTHVRCVILENDVQCDLDLYK SIQAKDHLKHTDHEITLONKILKREVRGHLKKRPEECDCCHRIYSHQHKGRLL KSSLHPKGLQEKDKWLLREQKXKLLKLLKLNQNDTCRWGLTCFTFDNHHW QATWLIGPFCACSTANNITNLTINEHNFLECFATGTFYFDFLSTDPQLMN ENKIDRDVNLQHLVQLMELRSKGYKQCNPNRNDLGLRDGGSYEQTRQRRKWP EMKRPSLSLGLQWEGWG"
BASE COUNT	951 a 1068 c 948 g 737 t
ORIGIN	
Query Match	62.3%; Score 2093.6; DB 11; Length 3704;
Best Local Similarity	86.1%; Pred. No. 0;
Matches 2360; Conservative	0; Mismatches 359; Indels 21; Gaps 3;
QY	1 GACTCCCGCATCCCAAAGACACAGATCAGCAAAAAGAGATGGCCCGCGAGC 60
Db	165 GGCCTCTGTCATCAGGAAGACAAAGACAGCAGCCCAATGSCACCCCTGSC 224
QY	61 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db	225 CTGCCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284
QY	121 CTGTCGCACACCGCTGAAAGCGAGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 180
Db	285 CTCTCCCATCCCGCTGAGGAGCGCTTCAGAGGACCGCAGGAACATCCGCCCAAC 344
QY	181 ATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db	345 ATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
QY	241 AAGACCGCGCATCATGAGGAGCGCGCGCGCATTCATCAACGCTTCGTGACCA 300
Db	405 AAGACCGCGCATCATGAGGAGCGCGCGCGCATTCATCAACGCTTCGTGACCA 464
QY	301 CCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db	465 CCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524
QY	361 ACTTACACCAATGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db	525 ACCTACACCAATGAGAAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584
QY	421 ACCTTTGCGGTGTACCTCAATAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 585 ACCTTCGCGGTGTATCTCAACAGCAGCAGCTACCGGACAGCTTTCTTCGGAAAAATACCTC 644  
Qy 481 AATGAATACAAACGGCTCTACTAGTCACACCGCGCTGGAAAGAGTGGGTGCGACTCTCTTAAA 540  
Db 645 AATGAGTACAAACGGCTCATAGTCGCGCGCGCGCTGGAAAGAGTGGGTGCGGCTTACTTAA 704  
Qy 541 AATCCCGCTTTTATACTACACGCTGTGTGCGAAACGGGTGAAAGAAAAAGCAGCGCTCC 600  
Db 705 AATCCCGCTTTTATACTACACACTCTGCGCGAATGGGTGAAAGAGAAACATGGGTCA 764  
Qy 601 GACTACTCAAGAGATTACTCAAGACCTCATACCAATGACACGCTGAGCTTCTTCGCG 660  
Db 765 GACTACTCAAGAGATTACTCAAGATCTCATACCAATGACACGCTGAGCTTCTTCGGA 824  
Qy 661 AGTCCTCAAGAGATTACTCAAGACCTCATACCAATGACACGCTGAGCTTCTTCGCG 720  
Db 825 ACATCAAGAGATTACTCAAGACCTCATACCAATGACACGCTGAGCTTCTTCGGA 884  
Qy 721 CACGCGCTGAGGATTACGCGCGCGCGCTCATACCAATGACACGCTGAGCTTCTTCGCG 780  
Db 885 CATGCGCGCGGAGCTACGCGCGCGCGCTCATACCAATGACACGCTGAGCTTCTTCGCG 944  
Qy 781 ATCAGCGCGAGCTACATACGCGCGCGCGCTCATACCAATGACACGCTGAGCTTCTTCGCG 840  
Db 945 ATCAGCGCGAGCTACATACGCGCGCGCGCTCATACCAATGACACGCTGAGCTTCTTCGCG 1004  
Qy 841 GGGCGCATGAAGCCCATCCATGGAATTCACCAATGTCACGCGGAGCGCTTTCAG 900  
Db 1005 GGACCCATGAAGCCCATTCATGGAATTCACCAATGTCACGCGGAGCGCTTTCAG 1064  
Qy 901 ACCCTCATGTCGCGGAGCTCATACGCGCGCGCTCATACCAATGTCACGCGGAGCGG 960  
Db 1065 ACCCTCATGTCGCGGAGCTCATACGCGCGCGCTCATACCAATGTCACGCGGAGCGG 1124  
Qy 961 GAGCTGGAACAACAGTACATCGTATACACCGCGGAGCGCTTTCAG 1020  
Db 1125 GAGCTGGAACAACAGTACATCGTATACACCGCGGAGCGCTTTCAG 1184  
Qy 1021 GGCCTGGTGAAGGGAATCATGCCATATGAGTTGACATCAGGTCGCGTCTACG 1080  
Db 1185 GGCCTGGTGAAGGGAATCATGCCATATGAGTTGACATCAGGTCGCGTCTACG 1244  
Qy 1081 AGGGGCCCCAAGTGAAGCGCGCTGTCTGAATCCCGAGCTGCTCTCAACATGACCTG 1140  
Db 1245 AGGGGCCCCAAGTGAAGCGCGCTGTCTGAATCCCGAGCTGCTCTCAACATGACCTG 1304  
Qy 1141 GCGCCACCATCTGGAATTCAGGCGCTGGAATTCAGGCGGAGTGAAGCGGAAATCC 1200  
Db 1305 GCGCCACCATCTGGAATTCAGGCGCTGGAATTCAGGCGGAGTGAAGCGGAAATCC 1364  
Qy 1201 ATCTCTAAGCTGCTGGAACAGCGCGCGCTGGAATTCAGGCGGAGTGAAGCGGAAATCC 1260  
Db 1365 ATCTCTAAGCTGCTGGAATTCAGGCGCGCGCTGGAATTCAGGCGGAGTGAAGCGGAAATCC 1424  
Qy 1261 AGGCTCTGGCGGAGTCTCTTCTGGTGAAGAGCGGAGCTGCTTACCAAGAGAGACAAT 1320  
Db 1425 AGGCTCTGGCGGAGTCTCTTCTGGTGAAGAGCGGAGCTGCTTACCAAGAGAGACAAT 1484  
Qy 1321 GACAGGTGAGCGCGCGGAGGAGTCTTCTGCGCGGAGTGAAGCGGAGTGAAGCGGAGT 1380  
Db 1485 GACAGGTGAGCGCGCGGAGGAGTCTTCTGCGCGGAGTGAAGCGGAGTGAAGCGGAGT 1544  
Qy 1381 TGTGAGCGTGTGATACAGACGCGCGTGTGAGAGCTGGGACAGAGTGGCGAGTGTG 1440  
Db 1545 TGTGAGCGTGTGATACAGACGCGCGTGTGAGAGCTGGGACAGAGTGGCGAGTGTG 1604  
Qy 1441 GAGGACCGCAAGCGGAGCTGAGCTGATGAGTGAAGCGCGCGCTGCGGCT----- 1493  
Db 1605 GAGGACCGCTTCTGGGAGCGCTGAGCTGCAAAATGTAAAGCGCGCTTGTGTCG 1664  
Qy 1494 --GGGCGGCGAGCGCGCTCTCAACCTCGTGGCGCGGAGTGTGAGAGAGCTTCTTCCAAA 2622  
Db 1665 GCGGCTGCGCAGCGCGCTCTCAACCTCGTGGCGCGGAGTGTGAGAGAGCTTCTTCCAAA 2801

Qy 1552 GCCTGCACTGTGACA-----GCGGGGACTACAAAGCTCAGCTCGCGCGGAGCGCGG 1602  
Db 1725 GCCTGCACTGTGACAAGTGTGCGGCTGAGGGGACTACAAAGTGGGCTGCTGCGAGCGCG- 1783  
Qy 1603 AAAAACTCTTCAAGAAGAGTACAAAGGCGCAGCTATGTCCGCGCTGCTTCCATCCGCTCA 1662  
Db 1784 --TAAGCTCTTTAAGAAAAAGTATTAAGACCAAGCTATGCGCGGACCGCTCCATCCGTTCC 1841  
Qy 1663 GTGGCCATCGAGGTGGAGCGGAGGTGTACCAAGTGGGCTGGGTGATGCCGCGCGAGCC 1722  
Db 1842 GTGGCCATCGAGGTGGAGCGGAGGTATACCAAGTGGGCTGGGTGATGCCGCGCGAGCC 1901  
Qy 1723 CGAAACCTCACCACGCGCACTGGGCGAGGGGCGCTGAGGACCAAGATGACAAAGAGTGGT 1782  
Db 1902 CGAAACCTTAGCAAGCGCACTGGGCGAGGGGCGCTGAGGACCAAGATGACAAAGAGTGGT 1961  
Qy 1783 GGGGACTTCAAGTGGGCACTGGGAGGCGCTTCCGAGCTACTCAGCGCGCGCAACCCATTAAGTG 1842  
Db 1962 GGCAGTTTCAAGTGGTGGTGGCTTCCAGATATTTCTGCGCGCGCAATCCCATCAAGTG 2021  
Qy 1843 ACACATCGGTGTACATCTAGAGAAACGACACAGTCCAGTGTGACCTGAGCTGTACAG 1902  
Db 2022 ACCCATCGGTGTACATCTTGAAGATGACACAGTCCAGTGGGACTTGGACCTGTACAG 2081  
Qy 1903 TCCCTGCGGCGCTGAAAGACCAAGCTGCAATTCGACCAAGAGATTTGAAACCTTGAG 1962  
Db 2082 TCCCTGCGGCGCTTGAAGAGACCAAGCTGCAATTCGACCAAGAGATTTGAAACCTTGAG 2141  
Qy 1963 AACAAATTAAGAACTCAGGGAGTCCGAGTGTGACCTGAGAAAGGCGGCGGAGAGAA 2022  
Db 2142 AACAAATTAAGAACTCAGGGAGTCCGAGTGTGACCTGAGAAAGGCGGCGGAGAGAA 2201  
Qy 2023 TGTGACTGTCAAAATCAGCTTACCAACCGACCAAGGCGCGCTCAAGCAGAGAGCG 2082  
Db 2202 TGTGACTGTCAAAATCAGCTTACCAACCGACCAAGGCGCGCTCAAGCAGAGAGCG 2261  
Qy 2083 TCCAGTGTGATCTCTTTCAGAGGGCGCTGAGAGAGAGAGAGAGTGTGCTGTGCGG 2142  
Db 2262 TCCAGCTGTGATCTCTTTCAGAGGGCTGAGAGAGAGAGAGAGTGTGCTGTGCGG 2321  
Qy 2143 GAGCAGAGAGCGCAGAGAAACTCCGCAAGCTGTCAAGCGCTGCAAGAAACCAAGAGAG 2202  
Db 2322 GAGCAGAGAGCGCAGAGAAACTCCGCAAGCTGTCAAGCGCTGCAAGAAACCAAGAGAG 2381  
Qy 2203 TGCAGATGCGCAGGCGCTCAGTGTCTTACCAACGAGCAACCGAGCTGGCAGAGCGGCT 2262  
Db 2382 TGCAGATGCGCAGGCGCTCAGTGTCTTACCAACGAGCAACCGAGCTGGCAGAGCGGCT 2441  
Qy 2263 TTTCTGACACTGGGCGCTTCTGTCGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 2322  
Db 2442 CTCTGAGCGCTGGGCGCGCTTCTGTCGCTGCAAGCGCGCGCGCGCGCGCGCGCGCG 2501  
Qy 2323 ATGAGGACCATCAATGAGACTCACAATTTCTTCTGTGAAATTTGCAACTTGGCTTCTTA 2382  
Db 2502 TTGAGGACCATCAATGAGACTCACAATTTCTTCTGTGAAATTTGCAACTTGGCTTCTTA 2561  
Qy 2383 GAGTACTTTGATCTCAACACAGACCGCTTACAGCTGATGATGAGTGAACACACTGAGAC 2442  
Db 2562 GAATACTTTGACCTCAGTACAGACCGCTTACAGCTGATGATGAGTGAACACACTGAGAC 2621  
Qy 2443 AGGGATGCTCTCAACAGCTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGCTTAC 2502  
Db 2622 AGGGAGCTCTTAAACCACTGCACTGCGAGCTCATGAGCTGAGGAGCTGTAAGGCTTAC 2681  
Qy 2503 AAGCAGTGTAAACCGCGAGCTGAAACATGAGACCTTGGACTTTAAAGAGAGCTAT 2562  
Db 2682 AAGCAGTGTAAACCGCGAGCTGAAACATGAGACCTTGGAGCTTTAGAGAGAGAGCTAT 2741  
Qy 2563 GAGCAATACAGCAGTGTTCAGCGTGAAGGTGGCGAGAAATGAAAGAGAGCTTCTTCCAAA 2622  
Db 2742 GACCAATACAGCAGTGTTCAGCGTGAAGGTGGCGAGAAATGAAAGAGAGCTTCTTCCAAA 2801

QY 2623 TCACTGGGCAACTGTGGGAAGCTGGGAAGTTAAGAAACACAGAGGTGGAGCTCCAA 2682  
 |||||  
 Db 2802 TCACTGGGACAGCTATGGGAAGCTTGGGAAGCTAAGCGGCATAGAGAGGAACCTCC 2861  
 |||||  
 QY 2683 AAACATAGAGGATCATCTGACTGACAGGCAATGAATA 2722  
 |||||  
 Db 2862 AAACACAGGGGCTCTGTGGCTGCCAGGCAATGAATA 2901  
 |||||

RESULT 2  
 LOCUS AK034712  
 DEFINITION Mus musculus 12 days embryo embryonic body between diaphragm region  
 and neck cDNA, RIKEN full-length enriched library, clone:19430027K19  
 product:weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix  
 coturnix], full insert sequence.

ACCESSION AK034712  
 VERSION AK034712.1 GI:26330130  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE 1  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 92279253  
 PUBMED 10349636

REFERENCE 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
 Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
 Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilmink, L.,  
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S.  
 and Hayashizaki, Y.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660

11217851

PUBMED  
REFERENCE  
AUTHORS

## TITLE

JOURNAL  
REFERENCE  
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 3759)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, P., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.

## Direct Submission

## TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

## source

1. 3759  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:9430027K19"  
 /db\_xref="taxon:10090"  
 /clone="9430027K19"  
 /tissue\_type="embryonic body between diaphragm region and  
 neck"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="12 days embryo"  
 160. 2886  
 /notes="unnamed protein product; putative  
 weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE  
 [Coturnix coturnix] (SPTR|Q90XB6, evidence: FASTY,  
 64.2%ID, 99.5%length, match=2601)"  
 /codon\_start=1  
 /protein\_id="BAC28804.1"  
 /db\_xref="GI:26330131"  
 /translation="VANPEGRGEIPSSRHHGLCTQBEAKDQQTAPPGLPLWLL  
 IVEQGAHFNAFTVTCPCPSRSITLGKYVHNHTYNNENSCSSPQWQAEHRTF  
 AYLNTSGVTAPEKGLYNEVSGYPPGKEMVGLKNRFYNTLCENGKEXKHS  
 DYSTDLTLIINDSVSPFRTSKMPHPHPVPLVISHAHPGPSAPQYSLFPNAS  
 QHTPSYNAIPNDRHINRYTGPMKPHMEFTNMLQRKLQTLMSVDSDMETIYDL  
 VETGLDNTIYLTADGHIQGLVKGKSMPEFDIRVFPYVGRPNVEAGSLNPHI  
 VLNIADPTLIYLTADGHIQGLVKGKSMPEFDIRVFPYVGRPNVEAGSLNPHI  
 KLHKREGKQVNAENFLPKYQVNDLCQAEYQACBQLGKQCVDSAGTLLKH  
 KCKPMRFQGGGSRALSNLVPKQSSACSDSGGGDYKLGLGRKLKPKKYK  
 TSVARNRSIRSVIAEVDGRIYHVGDTVPQENLSKPHWPAPEDQDDKGGFSGTG  
 GLPDSAPNPIKVTHRCYILENDTVQCDLILYKSLQAWDKHLHDHEIETLQNKQ  
 LRVRHKKKKPECCDCHRSYTHSGHGRKXKSSLPFRKRGQDKQKWLRLKQK  
 RKCKLKKLLQNNDTCSMPGLTCTHDHNNHWQAPLWTLGPPFCACYSANNNTWCL  
 RTINEHNFLEFCEFTYFDLSTDPYQLMANVNTLDRDVLNQLVQLMELRSCGK  
 YKCNPRTRNMDLGLRDGSGVEYQFQRRKWPENKRPSSKSLGLQWEGWG"

## CDS

BASE COUNT

962 a 1086 c 957 g 754 t



## ORIGIN

[illegible]

Qy	1021	GGCTGGTGAAGCGGAAATCCATGCCATATGAGTTTGA	CATCAGGGTCCCGTTCTACGTG	1081
Db	1234	GGGCTGGTGAAGGCGCAAGTCTATGCGCGTATGAAT	TTGCATCATAGAGTCCCGTTCTACGTG	1293
Qy	1081	AGGGGCCCCAACGCTGGAAGCGCGCTGTCTGAA	TCCCCCAATCGTCTCAACATTGACCTG	1140
Db	1294	AGGGGCCCCAACGCTGGAAGCTGCTCTGAA	CCCCCAATTTGCTCTCAACATTGACCTG	1353
Qy	1141	GCGCCACCATCTCGACATTCGACGCTCGACAT	ACTCTCGGATATGACGCGGAAATCC	1200
Db	1354	GCGCCACCATCTCGATATTCGTGACATTCGACAT	CCCTCGACATGACGCGGAAATCT	1413
Qy	1201	ATCCTCAAGCTGTGGACACGGAAGCGCGCGGTGA	ATCGGTTTCACTTGAATAAAGAAAGATG	1260
Db	1414	ATTCTCAAACTACTGACTCAGAGCGCGCAGTGA	ACCGGTTCCACTTGAATAAAGAAAGCTG	1473
Qy	1261	AGGGTCTGGCGGACCTCTTCTTGTTGGAGAGAG	GGCAAGCTGCTACACAGAGAGACAAT	1320
Db	1474	AGGGTCTGGCGGACCTCTTCTTGTTGGAGAGAG	GGCAAACTGCTCCACAAGAGGGAGGGT	1533
Qy	1321	GACAAGTGCACCCGAGGAGGAGAACTTTTCTG	CCCAAGTACACAGCGTGTGAAGGACCTG	1380
Db	1534	GACAAAGTGAATGCCGAGGAGGAGAACTTCTG	CCCAAGTACACAGCGTGTGAAGGACCTG	1593
Qy	1381	TGTCAGGGTCTGAGTACAGACGGGGTGTGACAG	CTGGACAGNAGTGGCAGTGTGTG	1440
Db	1594	TGTCAGGAGCTGAGTACAGACAGCACTGCGAA	CAGCTGGGGCAGAAAGTGGCAGTGTGTG	1653
Qy	1441	GAGGACGCCACGCGGGAAGCTGAAGCTGCATA	AGTGCAAAGGGGCCCACTGCGGGT----	1493
Db	1654	GAGGACGCTTCTGGGACGCTGAAGCTGCACAA	ATGTAAAGGGCCCCATCGGTTTGGTGGC	1713
Qy	1494	--GGGCGGACAGCAGCCCTCTCCACACTCTG	CGCCAGTACTACGGGACAGGCGACGAG	1551
Db	1714	GGCGGTGGCAGCAGACCCCTCTCCAACTTGG	TGGCCAAAGTATGACGGCCAGAGACGAG	1773
Qy	1552	GCCTGACCTGTGACA-----GCGGGGACT	CAAGCTCAGCCTGGCCGGAAGCGCCG	1602
Db	1774	GCCTGACGCTGTGACAGTGGCGGTGGAGGG	AGTACAAACTGGGCTTGCTGGAAGCGCC-	1832
Qy	1603	AAAAAATCTTCAAGAGAGATACAGGCGAGCT	ATGTCGCGAGTCTGCTCAATCCGCTCA	1662
Db	1833	--TAGCTCTTTAAGAAAAGATATAGACCA	CGCTATGCCCCGAAACGCTTCATCGTCCC	1890
Qy	1663	GTGGCCATCAGGTTGGAACGGCAGGGTGTAC	CAAGTACGAGGCTGGTGATGCCGCCAGCCC	1722
Db	1891	GTGGCCATCAGGTTGGAACGGTGGATATACC	ACGTTAGGCTTGGAATCTGTGCCTCAGCCC	1950
Qy	1723	CGAATCTCACAAGCGGCACCTGGCCAGGG	CGCCCTGAGGACCAAGATGACAGATGGT	1782
Db	1951	CGCACTTTGCAAGCGGCACCTGGCCAGGG	CGCCCTGAAAGCAAGATGCAAGGATGGT	2010
Qy	1783	GGGGACTTCAGTGGCACTGGAGGCTTCCG	GACTACTCAGCGCCCAACCCCAATTAAAGTG	1842
Db	2011	GGCAGTTTCACTGGTGGCTTCCAGATTTAT	CTGCGCCCCCAATCCCATCAAAGTG	2070
Qy	1843	ACACATCGGTGCTACATCTGAGAGACGAC	ACAGTCCAGTGTGACCTCGAGCTGTACAG	1902
Db	2071	ACCATCGGTGCTACATCTTGAGAAATGAC	ACAGTCCAGTGGCACTTGGACTCTGTACAG	2130
Qy	1903	TCCCTGCGGCTGGAAGACCAACAGCTGCA	CATTCGACAGAGATGAAACCTTGAG	1962
Db	2131	TCCCTGCGGCTTGAAGAGACCAACAGCTG	CACATTCGACCATGAGATGAAACCTTGAG	2190
Qy	1963	AACAAAATTAAGAACTGAGGGAATCGAG	GTGACCTCTGAAGAAAAGCGGCAGAGAA	2022
Db	2191	AACAAAATTAAGAACTTCGAGAGTCA	AGGGTCACTTGAAGAAAGCGGCAGAGAA	2250
Qy	2023	TGTGACTGTCACAAATCAGCTTACACAC	CCGACAAAGGCGCGCTCAAGCACAGAGC	2082
Db	2251	TGTGACTGTCATAGAAATCAGTTTAC	CACAGCCAAACAAAGGCGGTCTCAAGCACAAAGGC	2310



```
source 1..3687
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:4732465J15"
/db_xref="taxon:10090"
/clone="4732465J15"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
189..2816
/notes="unnamed protein product; putative
weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE
[Corumix coturnix] (SPR[Q90XB6, evidence: FASTY,
54.2%id, 99.5%length, match=2601])"
/codon_start=1
/protein_id="BAC26165.1"
/db_xref="GI:26324822"
translation="MAPPGPLPLLLSTALLSLLAGSAPLHPRLLKGRFORDRNIRP
NIIIVLDDDDVDELGSQVNVKTRINEQGAHPINAFVTPMCCPSRSILITGKYVH
NHTYTNENCSSPSQWQHESRTPAYLNSGTGRAPFGKYLNEINGSIVPPHPRVLM
VGLKNSRFTNITLCRNGVKEKGSDDYDILDTLINDSVSFFRTSKMYPHRPMKEM
VISHAAHPGDESDAPQSRFFPNASQHITPSYANPNPKHMYRGTGPKLVKRMFT
NMLQRKHLTMSYDDSMETIYDMLVETGELDNTIYLDADHGYHITGPKLVKRMFT
YEDIRVYFVYRGNVAGSLNPHIPLNIDLAIDTIDLAIDTIDLAIDTIDLAIDTID
YEPVNRPHLKKLRVWDSFLVBERKLLHREKDKVNAQENFLPKYQKDLQCORAE
YOTACEOLGOKWOCVEDASGTLKHKCKGPMRGGGGSRALSNLPKIDGQSSSEACS
CDSGGGDYKLGLAGRERKLKKYKASYANRSIRSVAIHVDGEI THVGLDITVPPQN
LSKPHNGAPEDQDDKGGSGTGGDLYSAPNP I KTVHRCVILENGTVQDLDLYK
SQAWKHDKHLHIDHEITLQNKLNREVRHLKKRPEEDCHRI SYHSHKHRLKH
KGSLSHPFRKLOEKDKVLRQKRLKRLKRLQNNDTCSMFGLTCFTHDNHWH
QTAPLWTLPFPFCACTSANNTWCLRTINETHNFCEPATGPIEVFDLSIDTPQYLMN
AVNTLDRDLNOLHVLMEILRSCKYKQCNPRNMDLGLRDGGSYEQIRQRRKWP
EMKEPSSKSLGQLEWEGH"
BASE COUNT 946 a 1062 c 948 g 731 t
ORIGIN

Query Match 62.2%; Score 2092; DB 11; Length 3687;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 360; Indels 21; Gaps 3;

QY 1 GACTCCCGCATCCCAAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60
DB 144 GGCCTCTGTCATCCAGGAGAGCAAGAGACAGCAAGCCACGCAATGGCACCCCTTGC 203
QY 61 CTCGTGTGTCTTGTCTCGCACTGTGTCTCTCCTGTGTGTGTGTGTGTGTGTGTGT 120
DB 204 CTGCCATATGGCTGTCTCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 263
QY 121 CTGTGGCAACACCCCTGAAAGGAGGATTTTTCAGAGGACCGCAGGAACATCCGCCCAAC 180
DB 264 CTCTCCCATCCCGCCTCAAGGAGACGCTTCCAGAGGACCGCAGGAACATCCGCCCAAC 323
QY 181 ATCATCTGGTGTGACGAGCAGCAGATGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGAAC 240
DB 324 ATCATCTGGTGTGTACGGATGTACAGATGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGAAC 383
QY 241 AAGACCGGCGCATCATGGAGCAGGCGGGCGGCACTTCATCAACGCTTTCGTGACCA 300
DB 384 AAGACAGGGGTATCATGGAGCAGGCGGGCGGCACTTCATCAATGCTTCTGTGACTACA 443
QY 301 CCCATGTGTCCTCTCAGCTCTCTCATCTCTCATCTGGCAAGTACGTTCACACCAAC 360
DB 444 CCAATGTGCTGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 503
QY 361 ACCTACACCAACATGAGACTGTCTCTCGCCTCTCTGGCAGCAGCAGCAGCAGCAGCAGC 420
DB 504 ACCTACACCAACATGAGATTTGTCTCTGCTCTCTCTGGCAGCAGCAGCAGCAGCAGCAGC 563
QY 421 ACCTTTGCGGTGATCACTCAATAGACATGGCTACCGGACAGCTTTCTTCGGGAGATATCTT 480
DB 564 ACCTTCGCGGTGATCTCAACAGCAGCAGGCTACCGGACAGCTTTCTTCGGGAGATATCTT 623
```

1704	Db		GCCTGCACTGTGACAGTGGCGGTGAGGGGACTACAAACTGGGCCTGGGTGACGCGCG - 1762
1603	Qy	AAAAA	CTCTTCAAGAAGAGTACAAAGGCCAGCTATGTCCGAGTGCCTCCATCCGCTCA 1662
1763	Db	--TAA	GCTCTTTAAGAAAAAGTATAGGGCCAGCTATGCCCGAACCGTCTCATCCGTTC 1820
1663	Qy	GTGCC	ATCGAGTGGACGGCAGGGGTATACACGTAGAGCTGGGTGATGGCGGCCACGCC 1722
1821	Db	GTGCC	ATCGAGTGGACGGTGAATATACACGTAGGCTTGGATCTGTGCTCTCAGCCC 1880
1723	Qy	CGAAC	CTCCAGCGGCACTGGCCAGGGGCCCTGAGGACCCAGATGACAGGATGGT 1782
1881	Db	CGAAC	CTTACAGCGCGCACTGGCCAGGGGCCCTGAGACCCAGATGACAGGATGGT 1940
1783	Qy	GGGAC	TTCACTGGCACTGGAGGCTTCCGCACTACTCAGCGCCCAACCCATTAAGTG 1842
1941	Db	GGCAT	TTCACTGGTACTGGTGGCTTTCAGATTATTTCTGCCCAATCCCATCAAAGTG 2000
1843	Qy	ACAT	CGGTGCTACATCTTAGAGACGACACAGTGCAGTGTGACTGGACCTGTGTACAG 1902
2001	Db	ACCAT	CGGTGCTACATCTTTGAGAAATGGCAAGTCCAGTGCCTTGGACCTGTGTACAG 2060
1903	Qy	TCC	TGCAGGCTGGAAAGACCAACAGCTGCACATCGACCAAGAGATTGAAACCTCTGCAG 1962
2061	Db	TCC	TGCAGGCTTGGAAAGACCAACAGCTGCACATCGACCAAGAGATTGAAACCTCTGCAG 2120
1963	Qy	AACAAA	ATTAGAACCTGAGGGAGTCCGAGGTCACTTGAGAAAAAGCGGCCGACGAGAA 2022
2121	Db	AACAAA	ATTAGAACCTTTCAGAGAGTTCAGGGGTCACTTGAGAAAAAGCGGCCGACGAGAA 2180
2023	Qy	TGTCAC	TGTCAAAAATCACTACACCCAGCACAAGGCCCGCTCAAGACACAGAGGC 2082
2181	Db	TGTAC	TGCCATAGATCATGTATACACAGCGACACAAAGGCCGTCTCAGACCAAGGC 2240
2083	Qy	TCCAG	TCTGATCTTTCAGGAAGGCTGCAAGAGAGGACAAAGTGTGCTGTGTCGG 2142
2241	Db	TCCAG	CTGCACTTTCAGGAAGGCTCTGAGGAGAGGACAAAGTGTGCTGTGTCGG 2300
2143	Qy	GACAG	AAGCGCAGAGAACTCCGCAAGCTGCTCAGCGCTTCGAGCAACAGACACG 2202
2301	Db	GACAG	AAACGCAAGAAAACTTGCCTCAAGCTGCTCAACAGGCTGCAAGACAGATACG 2360
2203	Qy	TGCAG	CTGCCCTCAGTGCTTTCACCCACGACACCCAGCACTGGCAGACGGCGCT 2262
2361	Db	TGCAG	CTGCCCTCAGTGCTTTCACCCACGACACCCACTGGCAGACGGCGCCA 2420
2263	Qy	TTCT	GGACACTGGGGCTTTCTGTGCTGCAACGGGCCAACATTAACGTACTGTGTC 2322
2421	Db	CTCT	GGACGTGGGGCGTTCTGCGCTGCAACGGGCCAACAAACAGTACTGTGTC 2480
2323	Qy	ATGAG	GCATCAATGAGACTCAAAATTCCTCTCTGTGAATTTGCAACTGGCTTCCCTA 2382
2481	Db	TTGAG	GCATCAATGAGACCCACAACTTCTCTCTGCAATTTGCAACCGGCTTCAATA 2540
2383	Qy	GAGTA	CTTTGATCTCAACAGACCCCTTACAGCTGATGATGACAGTGAAACACTGGAC 2442
2541	Db	GAA	TACTTGAACCTCAGTACAGACCCCTCACCAGCTGATGAACCGGCTGAAACACTGGAC 2600
2443	Qy	AGGAG	TGCTTCAACAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTAC 2502
2601	Db	AGGAG	CTCTTAAACAACTGCACTGTGAGCTCATGGAGCTTAGGAGCTGTAAAGGCTAC 2660
2503	Qy	AAGCA	GTAAACCCCGGACTCGAAAACATGGACCTGGGACTTAAAGATGGAGGAGCTAT 2562
2661	Db	AAGCA	GTGCAACCCCGGACCCGCAATGGACCTGGGCTTTAGAGACGAGGAGGCTAT 2720
2563	Qy	GAGCA	TACAGGAGTTTCAGCGTTCGAAAGTGGCCAGAAATCAAGAGACCTTCTTCCAAA 2622
2721	Db	GAA	CAATACAGGCACTTTCAGCGTTCGAAATGGCCAGAAATGAAGAGACCTTCTTCCAAA 2780
2623	Qy	TCACT	GGGACAACTGTGGGAGGCTGGGAGGTTAAGAAACAAACAGAGGTCGACTTCCAA 2682

Db	2781	TCACCTGGGACAGCTATGGAAGGCTGGGAAGGCTAAGCGGCCCATAGAGAGAGAACTCC	2849
QY	2693	AAACATAGAGGCGATCACCTGCATTCGACAGCCCAATGAAAAA	2722
Db	2841	AAAACAGGGGCTCTGTGTGGCTGCCAGGCCATGCAAAA	2880
RESULT 4	AK081643	3734 bp mRNA linear	HTC 05-DEC-2002
LOCUS	Mus musculus	16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130058L07 product: weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert sequence.	
ACCESSION	AK081643		
VERSION	AK081643.1	GI:26349278	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadori, K., Matsuda, H., Aeshburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, D. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsch, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, J., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Koh-tsuki, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE	5		

**AUTHORS** The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL** Nature 420, 563-573 (2002)

**REFERENCE** 6 (bases 1 to 3734)

**AUTHORS** Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tgami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

# FEATURES

## source

1..3734

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM\_DB:Cl30058107"

/db\_xref="taxon:10090"

/clone="Cl30058107"

/tissue\_type="head"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stages="16 days embryo"

232..2859

/note="unnamed protein product; putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPTR|Q90XB6, evidence: FASTY, 64.2%ID, 99.5%length, match=2601)"

/codon\_start=1

/protein\_id="BAC38279.1"

/db\_xref="GI:26349279"

/translation="MAPPGLPLWLLSTALLSLLAGSAFLSHPLKGRFDRNRIRP NIIVLVDQDVEYGSQVWVNRTERMEQGAHFINAFVTPMCCPSRSILTGKVVH NNYTYNNESSSWQAQHEHSRFPVFNSTGRTRAPFFKLINEYNGSVYPPGKWE VGLKNRNYNTLCRNGKEKHSYDSTLYDLTINDSVSPFRSKYPPHVPVIM VISAAHPGSDSAPQYSRFPNASQHTPSYAPNPDKHWRNVTGPKTHMEFT NMLQRLQTLMSYDSDNETYILMVTGELDNTYILYADHGVHIGQFLVKGKMP YEFPIRVFVYRGVNVAGSLNPHI VLNIDLAFLIDAGLIPADMDGSLIKLIDS YEPVNRFLHKKLVRWDSFLVERGKLLHKEGDKVNAOENFLPKYQVLDICQRAE YQTACEQLQXWQCVEDASGTLKHKCKGPMRFGGGGSRALNLFKIDQGSSEAS CDSGGDYGKUGLAGRKLFFKRYKTSYARNRSIRVAIEVDGEIYHVGLDTPQPRN LSKPHWPGAPDQDGGSGFGTGLPDYSAPNPVKVTHRCVILENDTVQCDDLYK SLOWKDKHLIHIEITLQNKIKLREVRHLKKRPERCDCHRSYHSQKGRLLK XGSSLHPFRKLOEKVLLKREOKRKLKLLKRLQNDTCSMPGLTCTFHDNHHW QTAPLWTLGPPFCACTSANNTYWCILRTINETNLFCEFATGFIYEDLSTDFYQLMN AVNTLDRDLNOLHVOIWEIERSCKGYQCPNRTNMDLGLRDGSGSYEQRFQRFKWP EMKRPSSKSLQQLWEGNEG"

## CDS

Best Local Similarity 86.1%; Pred. No. 0; Matches 2359; Conservative 0; Mismatches 360; Indels 21; Gaps 3;

QY 1 GACTCCCGCATCCAAAAGAACAGCAGATCAGCAAAAAGAAAGATGGCCCCCGGAGC 60

DB 187 GGGCTCTGGATCAGGAAGAACAGGACAGCAAGCCACCCCAATGGCACTTGGC 246

QY 61 CTGGTCTCTGCTTGTGTCGCCAACTGTGTTCTCCCTCTCTGGTGGAGCTCGGCTTC 120

DB 247 CTGCCACTATGGCTGTGTCCACCGCTCTCCCTCTCCCTGTGGCTGGAGCTTCG 306

QY 121 CTGTGCAACACCGCTGAAAGGAGGTTTCAGAGGGACCGCAGGAACATCCGCCCAAC 180

DB 307 CTCTCCCATCCCGCTGAAGGAGCGCTTCCAGAGGAGCCGAGGAACATCCGCCCAAC 366

QY 181 ATCATCTCTGTGTGTCAGCGACACAGGATGTGGAGCTGGGTTCATGCAAGTGTATG 240

DB 367 ATCATCTCTGTGTGTCAGCGATGACAGGATGTGGAGCTGGGTTCATGCAAGTGTATG 426

QY 241 AAGACCCGGCGCATCATGGAGCAGGGCGGGCGGCACTTCATCAACGCCTTCGTGACCA 300

DB 427 AAGACAAGCGGTATCATGGAGCAGGGCGGGCGGCACTTCATCAATGCTTCGTGACTA 486

QY 301 CCAATGTGCTGCTCGCTCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 360

DB 487 CCAATGTGCTGCTCGCTCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 546

QY 361 AACTACCAACAATGAGAACTGTCTCTGCGCTCTCTGGAGGCAAGCAAGAGAGCGCC 420

DB 547 AACTACCAACAATGAGAACTGTCTCTGCGCTCTCTGGAGGCAAGCAAGAGAGCGCC 606

QY 421 AACTTTGCGCTGTACTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCT 480

DB 607 AACTTTGCGCTGTACTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCT 666

QY 481 AATGAATACAAAGCTCTCTACAGCTCTCACTCACTCACTCACTCACTCACTCACTCAA 540

DB 667 AATGAGTACAACGGCTCTACTAGTCCGCGCGCGCTGGAAGGAGTGGTCTCGGCTACT 726

QY 541 AACTCCCGCTTTTAACTACAGCTGTGTGGAGCGGGGTGAAGAAGAAAGCAAGCGCTCC 600

DB 727 AACTCCCGCTTTTAACTACAGCTGTGTGGAGCGGGGTGAAGAAGAAAGCAAGCGCTCA 786

QY 601 GACTACTCCAAGGATTACTCTACAGCTCTCACTCACTCACTCACTCACTCACTCACTCAA 660

DB 787 GACTACTCCAAGGATTACTCTACAGCTCTCACTCACTCACTCACTCACTCACTCACTCAA 846

QY 661 AGTCCAGAAGATGTACCGGACAGGCGAGTCTCTATGGTTCATGAGCATGAGCGGCCCC 720

DB 847 ACATCCAAGAAGATGTACCGGACAGGCGAGTCTCTATGGTTCATGAGCATGAGCGGCCCC 906

QY 721 CACGGCCCTGAGGATTCAGCGCCCACTATTTACAGCTCTTCCCAAGCGCATCTCAGCAC 780

DB 907 CATGGCCCGGAGTCTAGCAAGCGGCTCTCAAGGCTCTTCCCAAGCGCATCTCAGCAC 966

QY 781 ATCAAGCGGAGTACAATCACTGCGGCCCAACCGGCAAAACTGGATCATCGCTTACACG 840

DB 967 ATCAAGCGGAGTACAATCACTGCGGCCCAACCGGCAAAACTGGATCATCGCTTACACG 1026

QY 841 GGGCCCATGAGCCCATCCATGGAATTCACCAACTGCTCCAGCGGAGGCGCTTCGAG 900

DB 1027 GACCCCATGAGCCCATCCATGGAATTCACCAACTGCTCCAGCGGAGGCGCTTCGAG 1086

QY 901 ACCCTCATGTCTGTGGAGTATCCATGGAGACGATTTACAACTGCTGTTTGAACGCGGC 960

DB 1087 ACCCTCATGTCTGTGGAGTATCCATGGAGACGATTTATGACATGCTGGTGGAGACGCGG 1146

QY 961 GAGCTGGACAACAGTATCATCGTATACACCGGCAAGCAAGTACCATCGGCGAGTTT 1020

DB 1147 GAGCTGGACAACAGTATCATCGTATACACCGGCAAGCAAGTACCATCGGCGAGTTT 1206

QY 1021 GGCCTGGTGAAGGGAATCCATGCGCATATGAGTTTGACATCAGGGTCCCGCTTCTACGTG 1080

BASE COUNT 956 a 1074 c 957 g 747 t

Query Match 62.2%; Score 2092; DB 11; Length 3734;







Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, K., Okazaki, Y., Gojibori, T., Bono, H., Kasakawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavani, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kaniya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Satoh, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 2685)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.  
Direct Submission  
Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 2685  
/organism="Mus musculus"  
/mol\_type="mRNA"

TITLE  
JOURNAL

COMMENT

FEATURES  
source

/strain="C57BL/6J"  
/db\_xref="PANTOM DB:C330011A01"  
/db\_xref="taxon:10090"  
/clone="C330011A01"  
/cell\_type="ES cells"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
c1.1816  
/note="unnamed protein product; putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPTX[Q90XB6, evidence: PASTY, 64.2%ID, 99.5%length, match=260])"  
/codon\_start=2  
/protein\_id="BAC33584.1"  
/db\_xref="GI:26333842"  
/translation="HMEFTNMLQRKRLTMSVDSMETIYDMLVETGELDNTLYLTADHGHIQGFGLVKGKSWPYEDIRVPYVGRPNVAGSLNPHIVLNDLAPTLIDIA GLDIPADWDGSLKLLDSEPVNRFHLKLLVRWDSFLVERGKLLHKGSDKYNQA BENFLPKYQVADLCQRAEYOTACEQLQKQVCVEDAGTLKHKCKGFMFGGGG RALSNLVPIQDQSEACSDSGGGDYGLGLAGRLFKKIKYTSYARNISRSVAI EVDGEIYHVLDTVPQPNLSKPHWGPEDDDKGSFSGTGLPYSPNFIKVT HRCYILENDTVQCDLDLYSLQAWKHKLHIDHEITLQNKIKLREVRGLKKRPE ECDCHRSYHSQKRLKHGSLHPRFKGLQKDEKQVLLREQKRLKRLKRLQNDTCSMPLGLCTFDHNDHMTAPLWTLGPPFCACTSANNVTYCLRTINETNLFCEP ATGTLEYDLSTDYQLMNAVNTLDRDVLNQLHVLMLRSCYKQKQCNPRTRNDLG LRDGGSYQYRQPRKRWPEMKRPFSSKSLGLGWEGWEG"

BASE COUNT  
ORIGIN

Query Match 42.5%; Score 1430.8; DB 11; Length 2685;  
Best Local Similarity 86.3%; Pred. No. 1.5e-285;  
Matches 1624; Conservative 0; Mismatches 237; Indels 21; Gaps 3;  
QY 859 CACATGGAAATTCACCAACATGCTCAGCGGAGCGCTTGCAGACCTCATGTCGGTGGAC 918  
DB 2 CACATGGAAATTCACCAACATGCTCAGCGGAGCGCTTGCAGACCTCATGTCGGTGGAC 61  
QY 919 GACTCCATGAGAGCATTTACAACATGCTGCTGAGCGGCGAGCTGGACAACACGTCAC 978  
DB 62 GACTCCATGAGAGCATTTACAACATGCTGCTGAGCGGCGAGCTGGACAACACGTCAC 121  
QY 979 ATCGTATACCGCGGACACCGGTTACACATCGCGCCAGTTTGGCTGGTGAAGGAGAA 1038  
DB 122 ATCTGTGACACCGCGGACACCGGTTACACATCGCGCCAGTTTGGCTGGTGAAGGAGAA 181  
QY 1039 TCCATGCCATATGAGTTGACATCAGGCTCCGCTTCTACGTGAGGGGCCCCAACAGTGGAA 1098  
DB 182 TCTATGCGGTATGAATTCGACATCAGAGTCCCGTTCTACGTGAGGGGCCCCAACAGTGGAA 241  
QY 1099 GCGGCTCTCTGAATCCCAACATCGTCTCAACATGACCTGGGCCCCCACCCTCTCTGGAC 1158  
DB 242 GCTGGCTCTCTGAACCCCAACATGCTCTCAACATGACCTGGGCCCCCACCCTCTCTGGAT 301  
QY 1159 ATTGAGCGCTGGACATACCTCGCGGATATGACGCGGAAATCCATCTCAAGCTCTGGAC 1218  
DB 302 ATGCGTGGAGTGGACATCTCCCTGCAGACATGACGCGGAGTCTATTCTCAAACTACTGGAC 361  
QY 1219 ACAGAGCGCGCGGTGAATCGGTTTTCACCTTGAAGAAGAGATGAGGGTCTGGCGGACTCC 1278  
DB 362 TCAGAGCGCGCGGTGAATCGGTTTTCACCTTGAAGAAGAGCTGAGGGTCTGGCGGACTCC 421  
QY 1279 TTCTTGGTGGAGAGGCAAGCTCTCTACACAAGAGAGACAATGACAAGGTGGAGCGCCAG 1338  
DB 422 TTCTTGGTGGAGAGGCAAGCTCTCTACACAAGAGAGAGAGTGAAGAAGTGAATGATCCAG 481  
QY 1339 GAGGAGACTTTCTGCCCAAGTACAGCGGTGAGGACCTGTGTGACCGTCTGAGTAC 1398  
DB 482 GAGGAGAACTTCTGCCCAAGTACAGCGGTGAGGACCTGTGTGACCGTCTGAGTAC 541  
QY 1399 CAGAGCGGCTGTGAGAGCTGGGACAGAGTGGCAGTGTGTGGAGGACGCCACACGGGAG 1458  
DB 542 CAGAGCAGCATCGAAGAGCTGGGACAGAGTGGCAGTGTGTGGAGGACGCTTCTGGGACG 601  
QY 1459 CTGAAGCTGCATTAAGTGAAGGGGCCCCCATGCGGCT-----GGCGGCGAGCAGAGCC 1509

Db 602 CTGAAGCTGCACAAATGTAAGAGCCCAATGCGTTTGGTGGCGGTGGGAGCAGAGCC 661  
Qy 1510 CTTCTCCAACTCGTCCCAAGTACTACGGGAGGCGAGGAGGCTGCACTGTGACA -- 1567  
Db 662 CTCTCAACTCTGTGCTCCCAAGTATACGGCCAGAGACGAGGAGGCTGCACTGTGACAGT 721  
Qy 1568 -----GGGGGACTACAACTACAGCTCGAGCTGGCGGAGCGCCGAAAGAACTCTTCAAGAAG 1620  
Db 722 GCGGCTGGAGGGGACTACAACTGGCGCTGGCTGAGCAGCG --TAAGCTCTTTAAGAAA 778  
Qy 1621 AGTACAGGCGCAGTATGTTCGAGTGTGCTTCATCCCTCAGTGGCCATCAGGTGGAC 1680  
Db 779 AGTATAAGACAGCTATGCTCCGGAACCGCTCCATCCGTTCGTGGCCATCAGGTGGAC 838  
Qy 1681 GGCAGGGTGTACCAAGTAGGCTGGGTGATGCGGCCAGCGCCGAAACCTCACCAAGCGG 1740  
Db 839 GGTGAGATATACCAAGTAGGCTGGATAGTGTGCTCAGCCCGCAACTTACCAAGCGG 898  
Qy 1741 CACTGCGCAGGCGCCCTGAGGACCAAGATGACAGGATGTGGGACTTCAGTGGCACT 1800  
Db 899 CACTGCGCAGGCGCCCTGAGGACCAAGATGACAGGATGTGGGACTTCAGTGGTACT 958  
Qy 1801 GGAGGCTTCCGACTACTACGCGCCCAACCCATTAAAGTGACACATCGTGTACATC 1860  
Db 959 GGTGGCTTCCAGATATTTCTGCCCCCAATCCATCAAGATGACCCATCGTGTACATC 1018  
Qy 1861 CTAGAGACGACACAGTCCAGTGTGACTGTGACCTGTGTACAAGTCCCTCAGGCTGGAAA 1920  
Db 1019 CTTGAGAATGACACAGTCCAGTGGACTTGGACCTGTACAAGTCCCTCAGGCTTGGAAA 1078  
Qy 1921 GACCAAGCTGCATCGACACCAAGATGGAACCTTGCAGACAAATTAAGAACTG 1980  
Db 1079 GACCAAGCTGCATCGACATGACCTGAGATCGAATCCCTGCAAGAAATTAAGAACTT 1138  
Qy 1981 AGGGAAGTCCGAGTCACTGTAAGAAAAGCGCGCCAGAGGAATGTGACTGTCACAAAATC 2040  
Db 1139 CGAGAAGTCCAGGGTCACTGTAAGAAAGAGCGACCGGAAGAATGTGACTGCCATAGAATC 1198  
Qy 2041 AGCTACCAACCCAGCACAAGGCGCTCAAGCAGACAGGCTCCAGTCTGATCCCTTC 2100  
Db 1199 AGTTACCAAGCAACAAAGGCGCTTCAAGCAAAAGGCTCCAGCTGACCCCTTC 1258  
Qy 2101 AGGAAGGCGCTCAAGAGAGACAGAGTGTGGCTGTGGCGGAGCAGAAGCGAAGAG 2160  
Db 1259 AGGAAGGCTCGAGGAGAGACAGTGTGGCTGTGGCGGAGCAGAAGCGAAGAG 1318  
Qy 2161 AAATCCGCAAGTGTCTAAGCGCTCGAGAAACAAAGACAGTGTGAGCATGCGAGGCTC 2220  
Db 1319 AAATCGCAAGTGTCTAAGCGCTCGAGAAACAAAGACAGTGTGAGCATGCGAGGCTC 1378  
Qy 2221 AGTGCTTCAACCAAGCAGACAGTGTGGCAGCGGCGCTTTCTGACACTGGGGCT 2280  
Db 1379 AGTGCTTCAACCAAGCAGACAGTGTGGCAGCGGCGCTTTCTGACACTGGGGCT 1438  
Qy 2281 TTCTGTGCTGCAACGAGCGCCCAATACAGTGTGCTGTGATGAGCAGCACCATCAATGAG 2340  
Db 1439 TTCTGTGCTGCAACGAGCGCCCAATACAGTGTGCTGTGATGAGCAGCACCATCAATGAG 1498  
Qy 2341 ACTCAAAATTTCTCTCTGTAATTTGCACTGCTTCTGATGAGTACTTTGATCTCAAC 2400  
Db 1499 ACCCAATTTCTCTCTGTAATTTGCACTGCTTCTGATGAGTACTTTGATCTCAAC 1558  
Qy 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGGACAGGAGTGTCTCAACACAG 2460  
Db 1559 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGGACAGGAGTGTCTCAACACAG 1618  
Qy 2461 CTACAGTACAGCTCATGAGCTGAGGAGTCAAGGTTTCAAGCAGTGTAAACCCCGG 2520  
Db 1619 CTGACAGTCACTCATGAGCTTAAGGAGCTTAAGGCTCAAGCAGTGAACCCCGG 1678  
Qy 2521 ACTCGAAACATGAGCCTGGGACTTAAGATGAGGAGGAGTGTGACCAATACAGGAGTTT 2580

Db 1679 ACCGCAACATGCACTGGGCTTAGAGACGAGGAGGAGCTATGAACAATACAGGAGTTT 1738  
Qy 2581 CAGCGTCGAAGTGGCCAGAAATGAGAGACCTTCTTCAATCTACTGGGCAACTGTGG 2640  
Db 1739 CAGCGTCGAAGTGGCCAGAAATGAGAGACCTTCTTCAATCTACTGGGCAAGCTATGG 1798  
Qy 2641 GAAGGCTGGGAAGTTTAAAGAAACAAGAGGTGGACCTTCAAAAAACATAGAGGCACTACC 2700  
Db 1799 GAAGGTTGGGAAGTTTAAAGGCTATAGAGAGAGGAACTCTCAAAACACAGGGGCTCGTG 1858  
Qy 2701 TGACTGCACAGGCAATGAATAA 2722  
Db 1859 TGCTGCCCAGGCAATGCAAAA 1880

AK045002 4564 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length  
enriched library, clone:B130018L10 product:SULFATASE FP homolog  
[Rattus norvegicus], full insert sequence.

AK045002  
AK045002.1 GI:26336990  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Warahiki, K.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, C., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohrsuiki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (5821), 585-590 (2001)  
21085660  
11217851  
5

## REFERENCE

## AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 4564)

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

## TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Funako Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/

## FEATURES

## source

Location/Qualifiers  
1. 4564  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:B130018L10"  
/db\_xref="taxon.10090"  
/clone="B130018L10"  
/tissue\_type="parthenogenote"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="9.5 days embryo"  
382. 2970

## CDS

/note="unnamed protein product; SULFATASE FP homolog [Rattus norvegicus] (SPTF|AAL71906, evidence: PASTY, 96.4%id, 100%length, match=2610)  
putative"  
/codon\_start=1  
/protein\_id="BAC32179.1"  
/db\_xref="GI:26336991"  
translations="MKYSLWALLAVLGTQLLSLCTVRSORFRGRLOOEKNIRPN IILVLDODVEIGSLQWNTKIMEQGANFTNAPVTPMCCPSRSRSMLTGKTVHN HNYVTNNENSSSWQAMREPRFVILANNIGTAFPTAFVTPMCCPSRSRSMLTGKTVHN GLIKNGRPNYVCRNGIKKEGDFYAKDFYFDLITNESINYPKMSKRMYPHPFIMV IISAAHPGPDSPAPQSKLYPNASQHTTSPYANPMKXIMQVMTPLPHEFTN VLQRKQLTMSVDSVERLYNNVLSGELDNTYIIYTADHGVIHQFGLVKGKMPY DFQIRVPPFRGSPVIFPQVPIVNLIDPLTILDIAGLSDPSVDGKSLKLDLE KPGNRPRTNKAKIMRDTLVERGFKRKESGKNIQSNHLPKRVKELCOQARY QTACEPQGNWQCIETDSGLRHKCKGSPDILLTVRQNRNLYSRGLHDKDECHRD

SGYRSRSORKNORQFLRNKGTPKYKPRFVHTROTSLRSVEFECEIYDINLEEELOV  
LPPSLAKHDEGHGEGTGHQAAAGDIENMLADNNAVGLPATVTVTHKCPILPNT  
IHCRRELYQSRARLKHSHLPKFAAAQVDSQLQKFKHRRRKRKRKRCQGGQSY  
NKGKVKQKQKAKHSHLPKFAAAQVDSQLQKFKHRRRKRKRKRCQGGQSY  
PGLCTFTDHHNQWTFAPFNILGSCACTSSNNNTYVCLRTVNTNTHFLFCFATGFL  
YFDMNTDPYQLTNTVHTVERSILNQLHIQLMELSCQGYKQCNPRKSLDTGAKEGN  
YDPHRA"

BASE COUNT 1282 a 1079 c 1150 g 1053 t  
ORIGIN  
Query, Match 31.6%; Score 1062.4; DB 11; Length 4564;  
Best Local Similarity 65.7%; Pred. No. 2.3e-209;  
Matches 1666; Conservative 0; Mismatches 811; Indels 57; Gaps 6;  
QY 97 CTGCTGGGTGGAAGCTCGGCCTTCTGTCGCGACACCGCCCTGAAAGGAGGAGTTTCAGAGG 156  
DB 430 CTGCTGGGAAGCCTGTGTTCACCGTTCGCTCCAGAGGTTCCGAGGAGGATACAGCAG 489  
QY 157 GACCGCAGGAATACCGCCCAACATCATCTGTGTGTGTCGCGAGCGAGCAGGATGTGGAG 216  
DB 490 GAACGAAAAACATCAGACCCCAATATTATCTTGTGTCTACTGACGACGAGGATGTGGAG 549  
QY 217 CTGSGTTCATGTCAGGTGATGAACAAGACCGCGCATCATGAGGAGCGGGCGGGCGGAC 276  
DB 550 CTGSGTTCCTTGTGAAGTCTGAACAAGAGAGATCATGGAACAGGGTGGGGCCACC 609  
QY 277 TTCATCAACGCCTTGTGTAACACACCCATGTGTGTCCTCCCTCAAGCTCTCTCATCTCACT 336  
DB 610 TTCACCAATGCTTGTGTGACACCGCCATGTGTGTCTCATCAAGCTCATCTCACTCACT 669  
QY 337 GGCAGTACGTCACACACCAACCAACCTACACCAACCAATGAGACTGCTCTCGCCCTCC 396  
DB 670 GGGAGTACGTCGATACCAACCAATGTCTACACCAACCAATGAGAACTGCTCTCTCTCTG 729  
QY 397 TGGCAGGACACGACGAGCGCGACCTTTGCGGTGTACCTCAATAGACTTGGCTACCGG 456  
DB 730 TGGCAGGCAATGACGAGCGCTCGACCTTGTGTGTATCTCAACACACCGGCTACAGA 789  
QY 457 ACAGTTTCTTCGGGAGATATCTTAATGATACAAACCGCTCTACGTCGACACCGGCTGG 516  
DB 790 ACAGCTTTTGTGAAAATACCTCAATGAATACAAATGAGAGTACATCCCTCTCTGATGG 849  
QY 517 AAGGAGTGGGTGCGACTCTTAAAACTCCCGCTTTTATACTACACGCTGTGTGCGAAC 576  
DB 850 CGAGATGGCTCGGATTAATCAGAAATCTCGTTCTTAATATACATCTGTTGTGCGAAC 909  
QY 577 GGGGTGAAGAAAAGCAGCGCTCCGACTACTTCAAGGATTAATCTTACAGACCTCATCACC 636  
DB 910 GGCATCAAGGAGAGCATGGATTGATTATGCAAGGATTACTTCAAGACTTAATCACT 969  
QY 637 AATCAGCGGTGAGCTTCTTCGCGACGTCGAGAGATGTACCGGACAGCGGAGTCCCTC 696  
DB 970 AACCAGAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1029  
QY 697 ATGTGTCATGACCCATGACGCGCCCGCCCGCTGAGGATTAATTCAGCCCAATATTCACGC 756  
DB 1030 ATGTGTCATGACCCATGACGACCAACATGAGGCTTGAAGACTCCGCGCCCAAGTTTCAAAA 1089  
QY 757 CTCTTCCCAAGGATCTCAGCAGCATCAGCGCGAGCTACAACTACGCGCCCAACCCCGAC 816  
DB 1090 CTGTACCCCAATGCTCTCCCAACACATAACCCAGCTATAACTATATGACCCCAACATGAC 1149  
QY 817 AAACACTGGATCATGCGCTACACGGGCGCCCATGAAGCCCATCAATGAATTCACCAAC 876  
DB 1150 AAAACACTGGATCATGCGCTACACGGGCGCCCATGTTGCTATCAATGAGATTCAACCAAC 1209  
QY 877 ATGTCTCAGCGGAGAGCGCTTGCAGACCCCTCATGTCGGTGGAGAGCTCCATGGAGAGCAATT 936  
DB 1210 GTCTCTCAGCGGAGAGAGCGCTTGCAGACTCTCTGATGCTGAGCTCTGTCTGAGAGTTG 1269  
QY 937 TACAACATGCTGGTGGAGAGCGGGGAGCTGGACAAACAGCTACATCGTATACACCGCGAC 996  
DB 1270 TACAACATGCTGGTGGAGAGCGGGGAGCTGGACAAACACTTACATCATCTACATCGCGAC 1329

QY 997 CACGTTACACATCGGCCAGTTTGGCTGTGAAAGGAAAATCCATGCCATATGAGTTT 1056  
DB |||||  
QY 1330 CACGGCTACACATCGGACATGTTGGGCTGTGCAAGGGGAAATCCATGCCCTAGACTTT 1389  
DB |||||  
QY 1057 GACATCAGGGTCCGCTTCTACGTGAGGGGCCCAAACCTGGAAGCCGCTGTCTGAATCCC 1116  
DB |||||  
QY 1390 GACATCCGTGTGCTCTTCTTCAATCGTGGTCAAGCATAGAAACAGGGTGGATGCCA 1449  
DB |||||  
QY 1117 CACATCGTCCCTCAACATTCAGCTCGCCCAACCTCTGGACATTCAGAGCCCTGGACATA 1176  
DB |||||  
QY 1450 CAGATGTTCTGACATCGACCTAGCTCTTACCATCTCTGATATTCCTGCTCTTACTCT 1509  
DB |||||  
QY 1177 CTTGGGATATGAGCGGAAATCCATCTCAAGTGTCTGACACGAGCGCCGGTGAAT 1236  
DB |||||  
QY 1510 CTTCTGATGTGGATGTAATCTGTCTCTCAACTCTAGACCTGGAAAGCCGGGTAA 1569  
DB |||||  
QY 1237 CGTCTTCACTTGAAGAAAGATGAGGGTCTGGGGGACTCTTCTTGTGGAGAGGC 1296  
DB |||||  
QY 1570 AGTCTTCGGAACAAGAGGCGCAAAATCTGGCGTGATACATCTCTAGTGGAAAGGG 1629  
DB |||||  
QY 1297 AAGCTGTCTACACAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCC 1356  
DB |||||  
QY 1630 AAATTTTACGGAAGAGGAGAAATCCGGCAAGAAATATCCACAGTCAACCACTTGCCC 1689  
DB |||||  
QY 1357 AAGTACGAGCTGTGAAGGACCTGTGTGACGCTGTGTGATACGAGCGGCTGTGAGCAG 1416  
DB |||||  
QY 1690 AAATATGAGAGGGTCAAGGAAATCTGTCCAGCAAGCCAGATACCAAGACAGCTGTGAGCAG 1749  
DB |||||  
QY 1417 CTGGGACAGAAAGTGGCAGTGTGTGAGGAGCCCAAGCGGGAAGCTGGAAGTGCATAAGTGC 1476  
DB |||||  
QY 1750 CTTGGGAGAACTGGCATGTGATCGAGGACATCTCGGCAAGCTCCGAATCCACAGTGT 1809  
DB |||||  
QY 1477 AAGGCCCA-----TCGGGTGGGGCGGAGAGAGCCCTCTCCAACTCTGTCGCCCAAG 1530  
DB |||||  
QY 1810 AAGGACCCAGCGACTCTCTCAGCGTCCGTGAGATGCAAGTGCACCAACCTCTACTCTCGCGGA 1869  
DB |||||  
QY 1531 TACTACGGGCGGAGCGAGCCCTGCACCTGTGACGCGGGGACTACAGCTCAGCCTG 1590  
DB |||||  
QY 1870 TTGCATGACAAAGACAAAGATGCCATTTGAGGACTCTGCTATCGTGGAGCAAGC 1929  
DB |||||  
QY 1591 GCCGACGCCGGAAGAACTCTTCAAGAG-----AAGTCAAGGCCGAGC 1635  
DB |||||  
QY 1930 CAGAGAAAGATCAAAAGCAGTTCTGTAGGAACAAGGAACACCAAGTATAAGCCGAGA 1989  
DB |||||  
QY 1636 TATGTCGAGTCTGCTCCATCGCTCAGTGGCCATCGAGTGGAGCGAGGGTGTACAC 1995  
DB |||||  
QY 1990 TTGTGTCACACCCGCGAGACTCGGTCTCTGTGTGAATTCGAAGGTGAAATATATGAC 2049  
DB |||||  
QY 1696 GTAGGCTGGGTGATCGCGCCGAG-----CCCCGAAACCTCACCAGCGGCAC 1743  
DB |||||  
QY 2050 ATCAACCTGGAAGAGAAAGAACTGCAAGTGTACCAACCAAGGAGCATTTGCCAGGCCAT 2109  
DB |||||  
QY 1744 TGGCAGGGGCCCCGTGAGGACCAAGATGACAGGATGTGGGACTTCAGTGGCACTGGA 1803  
DB |||||  
QY 2110 GATGAGGGTCAACAGGGCTTCATTGGCCACCAGGCTGCTGTGTGATCATCAGGAATGAG 2169  
DB |||||  
QY 1804 GGCCTTCCGACTACTCAGCCG-----CCAAACCCCATTAAGATGACACATCGG 1851  
DB |||||  
QY 2170 ATGTTGGGAGACAGCAACATGCTGTGGGGTTACCCGCACTGTCCGGGTGACACACAG 2229  
DB |||||  
QY 1852 TGCTACATCTTAGAAGACGACACATCCAGTGTGACCTGGACCTGTGTACAGTCCCTCGAG 1911  
DB |||||  
QY 2230 TGCTTTATTTCTTCCAAACGACACATCCACTGTGAGAGGAGCTGTACAGTCAAGCCAGA 2289  
DB |||||  
QY 1912 GCTCGAAGACCAAGCTGACATCCAGCAGATTTGAACCCCTGACAGACAAAT 1971  
DB |||||  
QY 2290 GCGTGAAGGACCAATAGGCCCTACATTTGATAAGAGATTTGAATTTCTCAAGATTAAT 2349  
DB |||||  
QY 1972 AAGAACTCTGAGGGAAGTCCGAGGTCACCTGTGAAGAAAAAGCCGAGAGAAATGTGACTGT 2031  
DB |||||  
QY 2350 AAGATTTAAGGGAAGTGAAGGGGACACCTAAAGAAAAAGAAACCTGAGGAGTGTGGCTGT 2409  
DB |||||

QY 2032 CACAAAATCAGTCTACCAACCCAGCACAAAGG-----CCGCTTCAAGCACAGAGGCTCC 2085  
DB |||||  
QY 2410 GGTGACACAGAGCTATTACAAACAGAGAAAGGTGTCAACACACAGGAGAACTAAAGAGT 2469  
DB |||||  
QY 2086 AGTCTGCATCTTTTCAGGAAGG-----CTGCAAGAGAAAGGCAAGGTGTGGCTGTG 2139  
DB |||||  
QY 2470 CACCTTTCACCCCTTTCAGGAGGCTGTGCTCCAGAGGTGGATAGCAAACTTCAGCTCTTC 2529  
DB |||||  
QY 2140 CGGAGAGCAGAGCGCAAGAACTCCGCAAGTGTCTCAAGCGCTCTGCAAAACAAACGAC 2199  
DB |||||  
QY 2530 AAGGAGCATCGAGGAGAGAGGAGGAGGAGAAACCGCAGAGGAGGAGAG 2589  
DB |||||  
QY 2200 AGTCTGCATGCTCAGGCTCACTGCTGTTACCCACGACAAACCGACACTGTCGACAGCGCG 2259  
DB |||||  
QY 2590 GAGTGTAGCTGCTGCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTGCTGCTGCT 2649  
DB |||||  
QY 2260 CTTTCTGCACTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2319  
DB |||||  
QY 2650 CATTCTGGAATTTGGATCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2709  
DB |||||  
QY 2320 TGCATGAGGACCATCAATGAGACTCAAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2379  
DB |||||  
QY 2710 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2769  
DB |||||  
QY 2380 CTAGAGTACTTGTATCTCAACACAGACCCCTTACAGCTGATGATGATGATGATGATGATGATGATGAT 2439  
DB |||||  
QY 2770 CTGGAATATTTGACATGAATACGCTGCTTATCAGCTCAAAATACAGTACACACAGTA 2829  
DB |||||  
QY 2440 GACAGGATGTCTCAACAGCTACACATGAGCTGCTGAGCTGAGGAGCTGCAAGGCT 2499  
DB |||||  
QY 2830 GAAAGGAGCATCTTAAATCAGCTACACATGAGCTGCTGAGCTCCGAGCTGCCAAGGG 2889  
DB |||||  
QY 2500 TACAAGCTGTAAACCCCGGACTCGAAACATGAGCTGCTGAGCTTAAAGATGGAGGAGC 2559  
DB |||||  
QY 2890 TATAAACAGTGAACCCCAAGACCCCAAGAGCTGGAATTTGAGCTTAAAGAGGAGGAAAC 2949  
DB |||||  
QY 2560 TATGAGCAATACAG 2573  
DB |||||  
QY 2950 TATGACCCGACAG 2963  
DB |||||

## RESULT 7

AK078756

LOCUS

DEFINITION

TITLE

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AK078756 4613 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 15 days embryo male testis cDNA, RIKEN full-length  
enriched library, clone:8030446E1.8 product:SULFATASE FP homolog  
[Rattus norvegicus], full insert sequence.

AK078756.1 GI:26098091

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

2049374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Kono, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,

Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuenli, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schrim, L.M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyoko-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Haegawa, Y., Kawaji, H., Kohseki, S. and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 4613)

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

**TITLE**  
**JOURNAL**

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.res.riken.go.jp/  
 URL: http://fantom.res.riken.go.jp/  
 Location/Qualifiers  
 1..4613  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:8030446E18"

**FEATURES**  
 source

/db\_xref="taxon:10090"  
 /clone="8030446E18"  
 /sex="male"  
 /tissue\_type="testis"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="15 days embryo"  
 misc\_feature  
 1..4613  
 /note="SULFATASE FP homolog [Rattus norvegicus]  
 (SPR|AAL71906, evidence: FASTY, 96.4%ID, 100%length,  
 match=2610)"

BASE COUNT 1321 a 1083 c 1136 g 1073 t  
 ORIGIN  
 Query Match 31.6%; Score 1062.4; DB 11; Length 4613;  
 Best Local Similarity 65.7%; Pred. NO. 2.3e-209;  
 Matches 1666; Conservative 0; Mismatches 811; Indels 57; Gaps 6;  
 QY 97 CTGCTGGGTGGGAAGCTTCGGCCCTTCCTGCGCACCAACCCCTGMAAGCGAGGTTTCAGAGG 156  
 DB 520 CTGCTGGGAAGCCTGTGTTCACCGTTCGGTCCAGAGGTTCCGAGGAAGATACAGCAG 579  
 QY 157 GACCGAGGAACATCCGCCCAACATATCTCTGCTGTCGAGCGACGACCAAGATGGAG 216  
 DB 580 GAACGAAAAAACATCAGACCCCAATATTATCTTGTGCTCACTGACGACCAAGATGGCAG 639  
 QY 217 CTGGGTTCATCGAGGTGATGAGAACCGCGCGCATCATGAGCAGCGGGGGGGCGCAC 276  
 DB 640 CTGGGTTCCTTGGCAAGTCAAGAACAGACGAGAAAGATCATGGAACAGGGTGGGGCCACC 699  
 QY 277 TTCTCAACAGCCCTTCGTGACCAACACCCATGCTGCTGCCCTCAGCTCTCTCCATCTCACT 336  
 DB 700 TTCACCAATGCTTTGTGACCAACCGCCATGCTGTCCATCAGCTCATCTCATCTCACT 759  
 QY 337 GGCAGTAGCTCCCAACACCAACACCTACACCAATGAGAACTGCTCTCGCCCTCC 396  
 DB 760 GGGAGTAGTGTGTCATAACCAATGTCTACACCAACCAATGAGAACTGCTCTCTCCCTCG 819  
 QY 397 TGGCAGGCACAGCAGCAGCGCGGACCTTTGCGGTGTACCTCAATAGCACTGGCTACCGG 456  
 DB 820 TGGCAGGCACATGACGAGCCTTGGACCTTTGCTGTATCTCAACACACCGGCTACAGA 879  
 QY 457 ACAGCTTTCTGGGAAGTATCTTAATGAATACAAAGCTCTTACGTCGACCGCCACCGGCTGG 516  
 DB 880 ACAGCCTTTTGGAAAAATACCTCAATGAATACAAATGCGAGCTACATCCCTCTCGATGG 939  
 QY 517 AAGGAGTGGTGGAGTCTCTTAAAAACTCCCGCTTTTAACTACTACGCTGTGCGGAAC 576  
 DB 940 CGAGTAATGCTCGGATTAATCAGAAATCTCTGTTCTATAATTAATTAATCTGTTTGGCAAC 999  
 QY 577 GGGGTGAAGAAAAGCAGCGCTCCGACTACTTCCAGGATTAACCTCAGAGACCTCATCACC 636  
 DB 1000 GGCATCAAGGAGAGCATGGATTTGATTATGCAAGGATTAATTCAGAGCTTAATCACT 1059  
 QY 637 AATGACGGGTGAGCTTTCTCCGACGCTCCAGAGAGTGTACCCGACAGCGCAGTCTC 696  
 DB 1060 AACGAGAGCATTAATTAATTTAAAAATGTCTAAGAGAAATGTATCCCCCATAGGCCCATATG 1119  
 QY 697 ATGCTCATCAGCCATGACGAGCGCCCGCCCGCTCAGGATTCAGGCGCCCAACAATATTCACGC 756  
 DB 1120 ATGCTGATCGCCATGACGAGCACCACATGGCCCTGAGAGCTCCGCCCCACAGTTTCAAAA 1179  
 QY 757 CTCTTCCCAAAAGCATCTCAGCAGATCATCGCGGAGCTTAACTAGCGGCCCAACCGCGAC 816  
 DB 1180 CTGTACCCCAATGCTCCCAACACATAACCCAGCTTAATTAATGACCCGAAACATGAGC 1239  
 QY 817 AAACACTGGATCTGCGGTACACGCGGGCCCATGAAGCCCATCCACATGGAATTCACCAAC 876  
 DB 1240 AAACACTGGATCTGCGGTACACGAGGGCCCATGTTGCTATCCACATGGAGTTTCCCAAC 1299  
 QY 877 ATGCTCAGCGGAAGCGCTTCAGACCCCTCATATGTCGGTGGAGCACTCCATGGAGACGATT 936  
 DB 1300 GTCTCTCCAGCGCAACGACTCGAGACTCTGATGTGCTGAGTTCAGTTGACGAGGTTG 1359



QY	937	TACAACTGCTGGTTGAGACGGCGAGCTGGACAACGTAATCATGTAATACACCGCCGAC	936
DB	1360	TACAACTGCTGGTGGAGAGCGGGAGCTGGACAACATTTATCATCTTACACTGCGGAC	1419
QY	997	CACGGTTACCACTCGGCCTGAGTTTGGCTGGTGAAGGGAAATCCATCCATATGAGTTT	1056
DB	1420	CACGGCTACCACTGGACAGTTTGGCTGGTCAAGGGAAATCCATGCCCTACGACTTT	1479
QY	1057	GACATCAGGCTCCGGTTCTACTGTAAGGGGCCCAACGTTGGAAGCGGCTGTCTGAATCCC	1116
DB	1480	GACATCCGCTGGCTCTTCTTCAITTCGTGGTCCAAAGCATAGAAACCAAGGCTCGATAGTCCCA	1539
QY	1117	CACATCGTCCCTCAACATTGACCTTGGCCCCCACCACATCTCTGCAATTCGAGGCTCGACATA	1176
DB	1540	CAGATTGTCTGAACTACGACCTAGCTCTCAACATTCGGAATTCGCTGGTCTTGACTCT	1599
QY	1177	CCTCGGATATGGACGGGAAATCCATCTCAAGTCTCTGGACAAGGAGCGGCCGTGAAT	1236
DB	1600	CCCTCTGATCTGGATGTGTAATCTGTCTCAAACTTCTAGACTCGGAAAGCGGGTAAC	1659
QY	1237	CGGTTTCACTTGAAGAAAGAGATGAGGGTCTGGCGGACTCCCTTCTTGGTGGAGAGGCG	1296
DB	1660	AGGTTTCCGA CAACAAGAAGGCCAAATCTGGCGTGATACATCTCTAGTGGAAAGAGG	1719
QY	1297	AAGCTGCTACACAAGAGAGACAATGACAAGTGGAGCGCCAGGAGAGAACTTTCTGCC	1356
DB	1720	AAATTTTACGGAAGAGGAGGAATCCGGCAAGAATATCCACAGTCAACCACTTCCCT	1779
QY	1357	AAGTACAGCGTGTGAAGAGACTGTGTACGCTGTGTAGTACAGACGGCTGTGTAGAGAG	1416
DB	1780	AAATATGAGAGGGTCAAGGAACCTGTGCACAGCAAGCCAGATACAGACAGCGCTGTGACAG	1839
QY	1417	CTGGGACAAGAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAACTGCATAAGTGC	1476
DB	1840	CCTGGCGAAGACTGGCAGTGCATCGAGAGACATCCGGCAGCTCCGAATCCACAGTGT	1899
QY	1477	AAGGCCCCCA-----TGCGCTGGCGGCGAGAGCCCTCTCCAACTCGTCCCAAG	1530
DB	1900	AAGGACCCAGCGACCTGTCTACGGTCCGTGAGAATGCACGCAACCTTACTCTCGGGA	1959
QY	1531	TACTACGGCAGGCGAGGCTGCACCTGTGACAGCGGGACTACAAGCTCAGCGCTG	1590
DB	1960	TGTCATGACAAGACAAAGAGTGCATTGTAGGCACTCTGGCTATCGTTCAGCAGAGC	2019
QY	1591	GCCGGA CGCCGGAAAAAATCTTTCAAGAAG-----AGTACAAGGCCAGC	1635
DB	2020	CAGAAAGAATCAAGCGAGTTCTTGAGAAACAAGGGAAACCAAAAGTATAAGCCAGA	2079
QY	1636	TATGTCGCCAGTCCCTCCCTCAGTGGCCATCGAGGTGGA CGGACAGGTTGACCA	1695
DB	2080	TTTGTGCACACCCGGCAGACTCGGTCCCTGTCTGTGTAATTCGAAGGTGNAATATATGAC	2139
QY	1696	GTAGGCTTGGGTGATCGCGCCAG-----CCCGGAAACCTCACAAGCGGAC	1743
DB	2140	ATCAACCTGGAAGAAGAAGAACTGCAAGTGTATACCACAAGGAGCATTGCCAGGGCCAT	2199
QY	1744	TGGCCAGGGGCCCTGTAGGACCAAGATGACAAGATGTGTGGGAATTCAGTGGCACTGGA	1803
DB	2200	GATGAGGCTTCCACAGGGCTTCAITGGCCACAGGCTGTCTGTGGTGAACATCAGGAATGAG	2259
QY	1804	GGCTTCCCGACTACTCAGCG-----CCAAACCCATTAAAGTCACACATCGG	1851
DB	2260	ATGTTGGCAGACAGCAACATGTGTGGGGTTACCCGCCACTGTCCGGGTGACACACAG	2319
QY	1852	TGCTACATCTTAGAAGACGACAGTCCAGTGTGACTTGACCTTGACCTGTACAAGTCCCTCGAG	1911
DB	2320	TGCTTTATCTTCTCAAAACGACACAATCCACTGTGAGAGGGAGCTGTACAAGTCAGCCAGA	2379
QY	1912	GCCTGGAAGACCAAGCTGCATCGACCGACGAGATTGAAACCTCCAGAACCAATTT	1971
DB	2380	GGTGGAGGAGCCATAGAGGCCCTCAITTGATGAAGATTTGAAGTTCTCAAGATATAATTT	2439
QY	1972	AAGAACCTGAGGGAGTCCGAGGTCACTCGAAGAAAAGCGGCCAGAGAAGATGTGACTGT	2031

Db	2440	AAAGAAATTTAAGGGAAGCTGAGGGGACACCTAAAGAAAGAAACCTGAGGAGTGTGGCTGT	2499
QY	2032	CACAAAATCAGCTTACCACACACCAGCACAAGG-----CGCCTCAAGCACAAGAGCTCC	2085
Db	2500	GGTGACCGAGGCTATTACAAAGAGGAAGGTGTCAACAGCAGGAGAGCTAAAGAGT	2559
QY	2086	AGTCTGCATCTTTTCAGGAAGGG-----CTGCAAGAGAAGGAAGAGTGTGGCTGTGTTG	2139
Db	2560	CACCTTCAGCCCTTCAAGAGAGGCTGTGTCGCCAGGAGTGTGATAGCAAACTTCAGCTCTTC	2619
QY	2140	CGGGAGCAGAGCGCAAGAGAACTCCGCAAGAGTCTCAAGCGCTGCAGAACCAACGAC	2199
Db	2620	AGGAGCATCGAGGAGGAGAAAGAGAGAGAGGAGAAAGCAGCAGAGGAGGAGAG	2679
QY	2200	AGTGAGCATGCGAGGCTCAAGTCTTACCCAGCAACACAGCACTGGCAGACGGCG	2259
Db	2680	GAGTGTAGCTGCTGCGCTTACCTGCTTACGCCATGACAACCACTGGCAGACTGCC	2739
QY	2260	CGTTCTGAGACATGGGGCTTTCTGTGCTCAGCAGCGCCAAACAATAACAGTACTGG	2319
Db	2740	CCATTCTGGAACCTGGGATCTTCTGTGCTGCTGCAAGATTCTAAACAATACCTACTGG	2799
QY	2320	TGCATGAGGACCATCAATGAGACTCAAAATTTCTCTTCTGTGAATTCGAATGCGCTTC	2379
Db	2800	TGTTTCGCTACAGTCAACGAGCGCACAATTTCTGTTTGTGAGTTTCTACTGGCTTT	2859
QY	2380	CTAGAGTACTTTGATCTCAACAGACAGCCCTACCAGCTGATGATGATGACGACACTG	2439
Db	2860	CTGGAATATTTCGACATGAATACGGATCTTATCAGCTCACAATACAGTACACACAGTA	2919
QY	2440	GACAGGGATGTCCTCAACAGGCTACACGTPACAGCTCATGAGAGCTGAGGAGCTGCAAGGT	2499
Db	2920	GACGAGGAGCATCTTAATCAGCTACACATACAGCTGATGAGCTCCGAGCTGCCAAGG	2979
QY	2500	TACAAGAGTGTAAACCCCGGAGCTCGAAACATGAGAGCTGGAGCTTAAAGATGAGGAGC	2559
Db	2980	TATAAACAGTGCACCAACCAAGAGCCAGAGCCTGGACATTTGGAGCTAAAGAGGAGAAAC	3039
QY	2560	TATGAGCAATACAG	2573
Db	3040	TATGACCCGACAG	3053
RESULT 8			
AK028285			
LOCUS		4428 bp	mRNA
DEFINITION		Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:323040B12 product:SULFATASE FP homolog [Rattus norvegicus], full insert sequence.	linear HTC 05-DEC-2002
ACCESSION			
VERSION		AK028285.1	GI:26390192
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE			
AUTHORS			
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			



Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multipillarary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861  
 4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Tomita, M., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilmberg, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851  
 5

Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 4428)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 Location/Qualifiers  
 1. 4428  
 /organism="Mus musculus"

## COMMENT

## FEATURES

source

/mol\_type="mrna"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:3230401B12"  
 /db\_xref="taxon:10090"  
 /clone="3230401B12"  
 /issue\_type="head"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="14, 17 days embryo"  
 287. - 2899  
 /note="unnamed protein product; SULFATASE FP homolog [Rattus norvegicus] (SPTR|AAL71906, evidence: PASTY, 96.4%ID, 100%length, match=2610)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC25858.1"  
 /db\_xref="GI:26390193"  
 /translation="MKYSLWALLAVLTGOLLGSLCTVRSORPRGRIOQRKNIRPN IILVTDODVELGSLQVNMKTKIMEGGATFTNAFTVTPMCCPSRSMLGKVVHN HNYVTNNCCSPSQMHEPRTPAVLNTGYRTAFPGKILNENYSGYIPFGWEWL GUKNSRFTNYTVCRNGIKERHGFYADYFTDLITNESINFFKYSKMSYIPRPMV ILSHAAPGDESAPOFSKLYPNASQHTPSYNAFMDKHWIMQYTGPMPLPHMSFTY VLQKRLQTLMSVDSVERLVNMLVSEGLDNTYIYTADHGYYHIGQFGLVKGSMPY DFDIRVPFFIRGSPISPEGSIVQIVLNTDLAPTILDIAGLDSPVKRVKELCOQARY KPNRFTNKKAKINRDTFLVERGKFLKESGKNIQSNHLPKRVKELCOQARY QTACQPGNWCIEDTSGKLEIKKCGPSDLLTVHONARLYSGLHDKCECHRD SGFRSRQKQKQRLNKGPKTKPRVTRQTRSLSEFEIGIYDINLEELQV LPPRSIAKHDEHQGFIGHQAAGDIRNEMLADSNNAVGLPATVTVTHKCELENDT ICEKRELYOSARAWDKHAYIDKEIVLQDKIKNLRVRHGLKCKPEKQCGGDSYY NKEGKROKLEKHLHPFKEAQAQVDSKLQLFKEHRRKRRKKEKRRKRGESCSL PGLTCTDHNNHWQTPAFNWSLFCACTSNNNTVCLRTVNETNFTLFCFEPATGFLE YFDMNTDYPQLNTVHTVRSILNQLHQLMELRSCQGVKQCNPRPKSLDIDGKGGN YDHPKQLWDGEG"

polyA\_signal 4414. 4419  
 /note="putative"  
 4428  
 polyA\_site /note="putative"  
 BASE COUNT 1278 a 1049 c 1077 g 1024 t  
 ORIGIN

Query Match 31.5%; Score 1060.8; DB 11; Length 4428;  
 Best Local Similarity 65.7%; Pred. No. 5e-209;  
 Matches 1665; Conservative 0; Mismatches 812; Indels 57; Gaps 6;

QY 97 CTGCTGGGTGGAGCTCGCCCTTCCTGTGCGACCCCGCTCGAAAGGCGAGGTTTCAGAGG 156  
 DB 335 CTGCTGGGAGCGCTGTGTCCACCGTTCGTCGCGAGAGGTTCCGAGAGGATACAGAG 394  
 QY 157 GACCGCAGGAACATCGCCCCCAACATCATCTCTGGTGTGACCGACGACGAGGATGTGGAG 216  
 DB 395 GAACGAAAAAATCATCAGACCCCAATATATCTGTGTCTCATCTGACGACCCAGGATGTGGAG 454  
 QY 217 CTGGGTTCATCAGGTGATGACAGCCCGCGCATCATGAGCAGGCGGCGGGCGGCAC 276  
 DB 455 CTGGGTTCCTCGAGTCATGAACAGACGAGAAAGATCATGGAACAGGTTGGGGCCACC 514  
 QY 277 TTTCATCAAGCCCTTCGTGACACACCCCATGTCTGCCCTCACGCTCTCTCCATCTCTCACT 336  
 DB 515 TTACCAATAGCTTTGTGACACGCGCCCATGTCTGTCCATCAGCTCATCTCATCTCACT 574  
 QY 337 GCGAAGTAGTCTCAACACCAACACCTTACCAACATGAGAACTGCTCTCGCCCTCC 396  
 DB 575 GGGAGTAGTGTGATCAACCAATGTCTACCAACAATGAGAACTGCTCTCGCTCTCG 634  
 QY 397 TGGCAGGCAGCAGCAGAGCGCGACCTTTGGCGGTGTACCTCAATAGACACTGGCTACCGG 456  
 DB 635 TGGCAGGCATGACGAGCGCTCGACCTTTGTGTGTATCTCAACAACACCGGCTACAGA 594  
 QY 457 ACAGCTTTCTTGGGAAGTATCTTAATGAATAACAAGGCTCTCTACGTGCGACCGGCTGG 516  
 DB 695 ACAGCTTTTGGAAAAATACCTCAATGAATAACATGAGGAGTACATCCCTCTCGATGG 754  
 QY 517 AAGGAGTGGTGGACTCCTTAAACCTCCCGCTTTTAACTACACGCTGTGTGCGAAC 576

755	CGAGNATGGCTCGGATTAAATCAGAATTTCTCGTTTCTATTAATTACACTGTTTGTGCGCAAC	814
577	GGGGTGAAGAAAAGCAGCGCTCCGACTACTCCAAAGATTACCTCACAGACTCATCACC	636
815	GGCATCAAGGAGAAGCATGGATTGATTATGATCAAAAGGATTACTTCACAGACTTAATCACT	874
637	AATCAGCAGGTGAGCTTCTTCCGACGTTCCAGAGATGTACCGCACAGCGCAGTCTCTC	696
875	AACGAGAGCATTAATTTACTTCAAATGTCTAAGAGATGTATCCCAATAGGCCCAATTATG	934
697	ATGTCATTCAGCCATGAGCCGCCCCACGCGCTTGAGGATTTCAGGCCCAATAATTTCACGC	756
935	ATGGTGATCAAGCCATGAGCACCACATGGCCCTGAAGACTCCGCGCCACAGTTTTCAAAA	994
757	CTCTTCCCAACAGCATCTCAGCACATCAGCGGAGCTACAATCAGCGCCCAACCGCGAC	816
995	CTGTACCCCAATGCTCTCCCAACACATAACACCCAGCTATAACTATTGCAACCGAATCGAC	1054
817	AAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATGGAATTCACCAAC	876
1055	AAACACTGGATCATGAGTACACAGGGCCCATGTTGCCATTCCACATGGAGTTCCACCAAC	1114
877	ATGCTCAGCGGAAGCGTTGCGAGACCTCATGTCTGGTGGACGACTCCCATGGAGACGATT	936
1115	GTCTCTCCAGCGCAAAAGCACTGCAGACTCTGATGTCAGTTGACGACTCTGTCTGAGAGGTG	1174
937	TACAAACATGCTGGTTGAGACGGCGAGCTGGACAACACAGCTACATCGTATACACCGCGAC	996
1175	TACAACATGCTTGTGGAGACGGGGAGCTGGACAACATTTACATCATCTACACTGGCGAC	1234
997	CACGGTTACCAACATCGCCACGTTTGGCTGGTGTGAAGGGAATCCCATCCATATGAGTTT	1056
1235	CACGGCTACCAACATGACAGATTGGCGCTGGTCAAGGGGAATCCATGCCCTACGACTTT	1294
1057	GACATCAGGTCCCGTTCTACGTGAGGGGCCCAAGCGTGGAGCCGGCTGTCTGAATCCC	1116
1295	GACATCCGTGTGCCCTTCTTCATTCTGGTGGTCCAGCATAGAACAGAGGTGCGATAGTCCC	1354
1117	CACATCGTCTTCAACATGACCTGGCCGCCCAACCACATCCTGGAATTCGACGGCTGGACATA	1176
1355	CAGATTGTTCTGAACATCGACCTAGCTCTCTACCATTTCTGATATTGCTGGTCTTGACTCT	1414
1177	CTGCGGATATGAGAGGGAATTCATCTCAAGCTGCTGGACACGAGAGCGGCGGTGAAT	1236
1415	CGTTCTGATGTGGATGTGTAATCTGTCTCTCAAACTTCTAGACTCGGAAAAGCCGGGTAAAC	1474
1237	CGGTTTTCATTTGAAAAGAGATGAGGTTCTGCGGGGACTCTCTTCTGTGTGAGAGAGGC	1296
1475	AGTTTTTCGACAAACAGAGAGGCCAAATCTGCGCTGATACATTTCTAGTGGAAAGAGGG	1534
1297	TAAGCTGTGTACACAGAGAGACAATGACAAGGTGGACGCCCAAGAGAGAACTTTTCTGCC	1356
1535	AAATTTTACGGAAGAGGAGGAATCCGGCAAGATATCCACAGTCAAAACCACTTGCC	1594
1357	AGTATACAGCTGTGAGGACCTGTGTACGCTGTGAGTGTACACAGCGCGCTGTGAGCAG	1416
1595	AAATATGAGAGGTTCAAGAACTGTGCACAGACCGAGATACACAGACCTGTGAGGAG	1654
1417	CTGGGACAGAAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGC	1476
1655	CCTGGGACAACTGGCAGTGCATCGAGGACACATCCGGCAAGCTCCGAATCCACAAGTGT	1714
1477	AAAGGCC-----CCATGGGCTGGGGCGGAGAGAGCGCTCTCCAA	1517
1715	AAAGGACCCACGACCTGCTCAGCGTCCATTCAGAAATGCAACGCAACTCTACTCTCGCGGA	1774
1518	CTCTGTGCCCAAG--TACTACGGGACGGGACGGAGCGCTGCACCTGTGACAGCGGGGAC	1575
1775	TTGCATGACAAAGACAAGAGTGCATTGTATGGGACTCTTGGCTATCGCTCGAGCAGAGC	1834
1576	TAACAAGTCAAGCTGGCGGACGCCGGAATAAACTCTTCAAGAGAGATTCACAGCCAGC	1635
1835	CAGAGAAAGAAATCAAAAGCAGTTCTCTGAGGAACAGGGAAACCAAGATTAAGCCAGA	1894

1636	Qy	TATGTCGGCAGTGCCTCCATCCGCTCAGTGGCCCATCGAGGTGACGCGCAGGTGTACCA	1695
1895	Db	TTTGTGCACACCCGGCAGACTCGGTCCCTGCTCTGTGTTAAATTCGAAGGTGAAATATATGAC	1954
1696	Qy	GTAGGCGTGGTGATCCGCCCAAG-----CCCGGAAACCTCACAAGCGGCAC	1743
1955	Db	ATCAACTTGGAAAGAAAGAACTGCAAGTGTATACCACCAAGGAGCATTCGCAAGCGCAT	2014
1744	Qy	TGGCCAGGGGCCCTTGAGAGACCAAGATGACAAGGATGTTGGGGACTTCAGTGCACCTGGA	1803
2015	Db	GATGAGGGTCAACAGGGCTTCATTTGGCCCAAGGCTGCTGCTGGTGCACATCAGGAATAG	2074
1804	Qy	GGCCTTCCCGACTACTCAGCG-----CCAAACCCATTAAAGTGAACACATCGG	1851
2075	Db	ATGTTGGCAGACAGCAACAATGCTGTGGGTTACCGCCACTGTCCTGGGTGACACACAAG	2134
1852	Qy	TGCTACATCTTAGAGAAAGACACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCCTGTCAG	1911
2135	Db	TGCTTTATTCTTCCAACAGCACAAATCCACTGTGTGAGGGAGCTGTACCAGTCAGCCAGA	2194
1912	Qy	GCCTGGAAGACCAACAAGCTGCACATCGACCAAGATTGAAACCCCTGCAGAACAAAAAT	1971
2195	Db	GCCTGGAAGGACCAATAGGCTTACATTGATAAAGAGATTGAAGTTCTACAGATATAAAT	2254
1972	Qy	AAGAACTGTAGGGAAGTCCGAGTGCCTGTAGAAAAAGCGGCCAGAAAGAAATGTGACTGT	2031
2255	Db	AAGAAATTAAGGGAAAGTGGGGACACCTTAAGAAAAAGGAAACCTGAGGAGTGTGGCTGT	2314
2032	Qy	CACAAAATCAGCTTACCACACCCACAGCAAAAG-----CCGCTCTCAAGCACACAGGCTCC	2085
2315	Db	GGTGCACGAGCTATTACACAAGAGAAAGGTGTCAAACGACAGGAGAGCTTAAAGAGT	2374
2086	Qy	AGTCTGCATCTTTACGAAAGG-----CCTGCAGAGAGAGCAAGGTGTGGCTGTTG	2139
2375	Db	CAGCTTCACCCCTTCAAGGAGGCTGCTGCTCCAGGAGGTGGATGACAAACTTCAGCTCTC	2434
2140	Qy	CGGAGCAGAAGCGCAGAAGAAACTCGCAAGCTGCTCAAGGCGCTGCAGAACACAACGAC	2199
2435	Db	AAGAGCATCGGAGGAGGAGAGGAGAGGAGAGAAAGAAACGGCAGAGGAGGAGAG	2494
2200	Qy	ACGTGCAGCATGCAGGCTTCAGTGCCTTCACCACGACAAACGACACTGGCAGACGGCG	2259
2495	Db	GAGTGTAGCTGCCTGGCTTACCTCTTCACCCATGACAAACCACTGGCAGACTGCC	2554
2260	Qy	CCCTTCTGGACACTGGGCTCTTCTGTGCTGCACAGCGCCAAACAATAACAGTACTGG	2319
2555	Db	CCATCTGGAACTTGGGATCTTCTGTGCTGCACAAGTTCTAACACAATACCTACTGG	2614
2320	Qy	TGCATGAGACCAATCAATGAGACTCAATTTCTCTCTCTGTGAAATTTGCACTGGCTTC	2379
2615	Db	TGTTTGGCTACAGTCAACGAGAGCGCAATTTCTCTTTTGTGAGTTGTCTACCTGGCTTT	2674
2380	Qy	CTAGAGTACTTTGATCTCAACACAGACCCCTACAGCTGATGATGAGTGAACACACACTG	2439
2675	Db	CTGGAATTTTCGACATGAATACGGATCCTTATCAGCTCACAAATACAGTACACACAGTA	2734
2440	Qy	GACAGGATGTCTCAACCAAGCTTACAGGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGT	2499
2735	Db	GAAACGGAGCATCTTAATCAGCTTACACATACAGCTGATGGAGCTCCGAAGCTCCCAAGG	2794
2500	Qy	TACAACGAGTGAACCCCGGACTCGAAACATGGACTGGGACTTAAAGATGAGGAAGC	2559
2795	Db	TATAAACAGTGCACCCAGACCCAGAGACCTGGACCTTGGAGCTTAAGAAAGGAGGAAC	2854
2560	Qy	TATGAGCAATACAG	2573
2855	Db	TATGACCCGACAG	2868

RESULT 9  
AK008108  
LOCUS

1803 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert sequence.

ACCESSION AK008108.1 GI:12842085

VERSION 1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Hill, D., Kackenberg, J., Schiraldi, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyono-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (5821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 1803)

AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATTCGAGTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

Location/Qualifiers

1..1803

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM,DB:2010004N24"

/db\_xref="MGI:1904769"

/db\_xref="taxon:10090"

/clone="2010004N24"

/sex="male"

/tissue type="small intestine"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

715..1164

/note="unnamed protein product; putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPTR|Q90XB6, evidence: FASTY, 64.2%ID, 99.5%length, match=2601)"

/codon\_start=1

/protein\_id="BAB25464.1"

/db\_xref="GI:12842086"

/db\_xref="MGI:1919293"

/translation="MPGLCTFDHNDHWQAPLWTLGFFCACTSANNTYVCLRTINE PRTNLFCEBATGFLEYPDLSDPYQLMNAVNTLDRDLNOLHVLQMLRSCKGYKQCN TRNFMGLGLRDGGSVEQYRQFRKWPENKRPSSKSLGLWEGWG"

1778..1783

/note="putative"

1803

/note="putative"

BASE COUNT 481 a 512 c 467 g 343 t

ORIGIN

Query Match 25.5%; Score 856.4; DB 11; Length 1803;

Best Local Similarity 87.0%; Pred. No. 1.1e-166;

Matches 941; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 1641 CCGCAGTCGCTCCATCCGCTCAGTGGCCATCGAGTGGAGCGGCGGTGTACACGCTAGG 1700

DB 147 CCGCGGCTCTCCATCCGCTCCGTCGTCATCGAGTGGAGCGGTGAGATATACACGCTAGG 206

QY 1701 CCGGTCGATCGCCGCCGCCGCCGCCAAACCTTACCACAGCGGCACTGGCGGCGGCCCTGA 1760

Db	207	CTTGATCTGCTGCTCAGCCCCCGACCTTAGACCGCAGCTGGCCAGGGGCCCTCTGA	266
Qy	1761	GGACCAAGATGACAAAGATGCTGGGAGCTTCAGTGGCACTGGAGGCTTCCCGACTACTC	1820
Db	267	AGACCAAGATGACAAAGATGCTGGGAGCTTCAGTGGCACTGGAGGCTTCCCGACTACTC	326
Qy	1821	AGCGCCAAACCCCTTAAGTGAACATCGTGTCTACATCTAGAGAAACACACAGTCCA	1880
Db	327	TGCCCCCAATCCATCAAGTGAACATCGTGTCTACATCTAGAGAAACACACAGTCCA	386
Qy	1881	GTGTGACCTGACCTGTACAAAGTCCCTGAGGCTGGAAAGACACAAAGTGGACATCGA	1940
Db	387	GTGCACTTGGACCTGTACAAAGTCCCTGAGGCTGGAAAGACACAAAGTGGACATCGA	446
Qy	1941	CCACGAGATTGAACCTTCGAGACAAATTAAGACCTGAGGAGCTCCGAGTCACTT	2000
Db	447	CCATGAGATCGAAACCTTCGAGACAAATTAAGACCTTCGAGAGTCAAGGGTCACTT	506
Qy	2001	GAAGAAAGCGGCGCAGAGAAATGTGACTGTACAAATCAGCTACACACCCAGCAAA	2060
Db	507	GAAGAAAGCGGCGCAGAGAAATGTGACTGTACAAATCAGCTACACACCCAGCAAA	566
Qy	2061	AGCGGCTCAAGCACAGAGGCTCCAGTCTGATCCCTTCAGGAGGCTTCGAGGAA	2120
Db	567	AGCGGCTCAAGCACAGAGGCTCCAGTCTGATCCCTTCAGGAGGCTTCGAGGAA	626
Qy	2121	GGACAAAGTGTGGCTGTTCGCGGAGCAGAGCGCAAGAAATCTCGCAAGCTGTCAA	2180
Db	627	GGACAAAGTGTGGCTGTTCGCGGAGCAGAGCGCAAGAAATCTCGCAAGCTGTCAA	686
Qy	2181	GGCGCTGAGAAACAGACAGCTGACATGCGAGGCTTCAGTGTCTACACACGACAA	2240
Db	687	AGCGGTGAGAAACAGACAGCTGACATGCGAGGCTTCAGTGTCTACACACGACAA	746
Qy	2241	CCAGCACTGGCAGACGGGCTTTCTGACACTGGGCTTTCTGCTGCTGACACGCGC	2300
Db	747	CCACCACTGGCAGACGGGCTTTCTGACACTGGGCTTTCTGCTGCTGACACGCGC	806
Qy	2301	CAACAAATGACCTGCTGCTGAGGACCATCAATGAGACTCACAAATTTCTTCTG	2360
Db	807	CAACAAATGACCTGCTGAGGACCATCAATGAGACTCACAAATTTCTTCTG	866
Qy	2361	TGAATTTGCACTGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTTACAGCTGAT	2420
Db	867	CGAATTTGCACTGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTTACAGCTGAT	926
Qy	2421	GAAATGAGTGAACACACTGGACAGGAGTGTCTCAACAGGCTACACGCTCATGGA	2480
Db	927	GAAATGAGTGAACACACTGGACAGGAGTGTCTCAACAGGCTACACGCTCATGGA	986
Qy	2481	GCTGAGGAGCTCAAGGTTACAGCAGTGTAAACCCCGGAGCTGAAACATGAGCTGGG	2540
Db	987	GCTGAGGAGCTCAAGGTTACAGCAGTGTAAACCCCGGAGCTGAAACATGAGCTGGG	1046
Qy	2541	ACTTAAAGATGAGGAGTATGAGCAATACAGGAGTTCAGGCTCGAAAGTGGCCAGA	2600
Db	1047	GCTTAAAGATGAGGAGTATGAGCAATACAGGAGTTCAGGCTCGAAAGTGGCCAGA	1106
Qy	2601	AATGAGAGACTTCTTCAAAATCACTGGGCAACTGTGGGAAAGCTGGGAGGTTAAGA	2660
Db	1107	AATGAGAGACTTCTTCAAAATCACTGGGCAACTGTGGGAAAGCTGGGAGGTTAAG	1166
Qy	2661	AACAACAGAGTGGACCTTCAAAATCACTAGAGGATCACTGATGACACAGGCAATGAA	2720
Db	1167	GCCCATAGAGAGGAGCACTTCAAAATCACTAGAGGATCACTGATGACACAGGCAATG	1226
Qy	2721	AA	2722
Db	1227	AA	1228

RESULT 10

AL537332	1201 bp	mRNA	linear	EST 31-MAY-2003
LOCUS				
DEFINITION	CSODF025YB15 5-PRIME, mRNA sequence.			
ACCESSION	AL537332			
VERSION	AL537332.2	GI:31261955		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
COMMENT	On Feb 13, 2001 this sequence version replaced gi:12800825.			
Genoscope - Centre National de Sequencage				
BP 191 91006 EVRY cedex - France				
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
Library was constructed by Life Technologies, a division of				
Invitrogen. This sequence belongs to sequence cluster 9923.r For				
more information about this cluster, see				
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODF025AC08QPl&cluster=9923.r. Contact :				
Feng Liang Email : fliang@lifetech.com URL :				
http://fulllength.invitrogen.com/Invitrogen Corporation 1600				
Paraday Avenue Genoscope sequence ID : CSODF025AC08QPl.				
Location/Qualifiers				
1. 1201				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="CSODF025YB15"				
/tissue_type="FETAL BRAIN"				
/dev_stage="fetal"				
/clone_lib="Homo sapiens FETAL BRAIN"				
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA				
was primed with a NotI-oligo(dT) primer. Five prime end				
enriched, double-strand cDNA was digested with Not I and				
cloned into the Not I and EcoRV sites of the pCMVSPORT 6				
vector. Library was not normalized."				
BASE COUNT	296 a	319 c	335 g	193 t
ORIGIN				
Query Match	24.9%	Score 838.2;	DB 9;	Length 1201;
Best Local Similarity	95.1%;	Pred. No. 6.5e-163;		
Matches	904;	Conservative	4;	Mismatches 37; Indels 6; Gaps 4;
Qy	1282	TTTGTGGAGAGAGGCAAGCTGTACACAAAGAGAGACAATGACAAAGTGGAGCGCCAGGAG	1341	
Db	20	TTTGTACAAAAGCAGGCTGTACCGGTCCGGAATCCCGGATGTGAGCCCGAGGAG	79	
Qy	1342	GAGAACTTTCTGCCAAGTACAGCGTGTGAAGGACCTGTGTACCGTGTGAGTACCAG	1401	
Db	80	GAGAACTTTCTGCCAAGTACAGCGTGTGAAGGACCTGTGTACCGTGTGAGTACCAG	139	
Qy	1402	ACGGCGTGTGACAGCTGGACAGAGTGGCAGTGTGTGGAGAGCGCCACCGGGAGGCTG	1461	
Db	140	ACGGCGTGTGACAGCTGGACAGAGTGGCAGTGTGTGGAGAGCGCCACCGGGAGGCTG	199	
Qy	1462	AAGCTGCATTAAGTGAAGGGCCCCATGCGGTGGCGGCGCAGAGCCCTCTTCCAACTC	1521	
Db	200	AAGCTGCATTAAGTGAAGGGCCCCATGCGGTGGCGGCGCAGAGCCCTCTTCCAACTC	259	
Qy	1522	GTGCCCCAAGTACTACGGGCGAGCGGAGGCGCTGACCTGTGACAGCGGGAGCTACAG	1581	
Db	260	GTGCCCCAAGTACTACGGGCGAGCGGAGGCGCTGACCTGTGACAGCGGGAGCTACAG	319	
Qy	1582	CTCAGCTGGCGGAGCGCGGAAAAAACTTTCAAGAGAAAGTACAAAGCCAGCTATGTC	1641	
Db	320	CTCAGCTGGCGGAGCGCGGAAAAAACTTTCAAGAGAAAGTACAAAGCCAGCTATGTC	379	

```
QY 1642 CGCAGTCGCTCCATCGCTCAGTGGCCATCGAGTGGAGCGGAGGGTGTACACCTAGGC 1701
Db 380 CGCAGTCGCTCCATCGCTCAGTGGCCATCGAGTGGAGCGGAGGGTGTACACCTAGGC 439
QY 1702 CTGGGTGATGCGCGCCAGCCGAAACCTCCACCAAGCGGCACTGGCCAGGGGCCCTCTGAG 1761
Db 440 CTGGGTGATGCGCGCCAGCCGAAACCTCCACCAAGCGGCACTGGCCAGGGGCCCTCTGAG 499
QY 1762 GACCAAGATGACAGATGGTGGGACCTTCACTGGGCACTGGAGGGCTTCCCGACTACTCA 1821
Db 500 GACCAAGATGACAGATGGTGGGACCTTCACTGGGCACTGGAGGGCTTCCCGACTACTCA 559
QY 1822 GCGCCCAACCCCAATTAAGTGCACATCGTGTCTACATCCTAGAGAACGACACATGCCAG 1881
Db 560 GCGCCCAACCCCAATTAAGTGCACATCGTGTCTACATCCTAGAGAACGACACATGCCAG 619
QY 1882 TGTGACCTGGACCTGTACAAAGTCCCTGAGGCTGGAAAGACCAACAGCTGCACATGCCAG 1941
Db 620 TGTGACCTGGACCTGTACAAAGTCCCTGAGGCTGGAAAGACCAACAGCTGCACATGCCAG 679
QY 1942 CACGAGATTGAACCTGACAGACAAATTAAGTGCACATCCTAGAGAACGACACATGCCAG 2001
Db 680 CACGAGATTGAACCTGACAGACAAATTAAGTGCACATCCTAGAGAACGACACATGCCAG 739
QY 2002 AAGAAAAGCGCCAGAGAAATGTGACTGTCAAAAATCAGTACACACCCAGCACAAA 2061
Db 740 AAGAAAAGCGCCAGAGAAATGTGACTGTCAAAAATCAGTACACACCCAGCACAAA 799
QY 2062 GCGCGCTCAAGCAGACAGGCTTCAAGTGTGATCTTTCAGGAGGGCTTCCAGAGAGAG 2121
Db 800 GCGCGCTCAAGCAGACAGGCTTCAAGTGTGATCTTTCAGGAGGGCTTCCAGAGAGAG 859
QY 2122 GACAAAGTGTGCTGTGCGGAGCAGAAAGCCAGAGAGAAATCCGCAAGCTGCTCAAG 2181
Db 860 GACAAAGTGTGCTGTGCGGAGCAGAAAGCCAGAGAGAAATCCGCAAGCTGCTCAAG 916
QY 2182 CCGCTGAGAACACGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2232
Db 917 CCGCTGAGAACACGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 964

RESULT 11
AL574439/c
LOCUS AL574439 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION clone CS0DI058YD10 3-PRIME, mRNA sequence.
AL574439
AL574439.2 GI:31312757
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12934646.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9923.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI058YD1058NP1cluster=9923.r. Contact :
Feng Liang Email : fliang@lifetech.com URL : Corporation 1500
http://fulllength.invitrogen.com/ Invitrogen
Paradise Avenue Genoscope sequence ID : CS0DI058YD1058NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI058YD10"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 273 a 253 g 334 t 100 others
ORIGIN
Query Match 24.7%; Score 829.4; DB 9; Length 1201;
Best Local Similarity 92.5%; Pred. No. 4.3e-161;
Matches 864; Conservative 40; Mismatches 21; Indels 9; Gaps 4;
QY 2391 TGATCTCAACACACAGACCCCTTACCAGCTGATGAATGAGTGAACACACACTGACAGGATGT 2450
Db 934 TTATYTCACACAGACCCCTTACCAGCTGATGAATGAGTGAACACACACTGACAGGATGT 875
QY 2451 CCTCAACCAAGTACACGTCAGCTCATGAGTGAAGAGCTGCAAGGGTTACAAGCAGTG 2510
Db 874 CCTCAACCAAGTACACGTCAGCTCATGAGTGAAGAGCTGCAAGGGTTACAAGCAGTG 815
QY 2511 TAACCCCGGACCTCGAAACATGAGACCTGAGCTTAAAGATGAGGAGCTATGAGCAATA 2570
Db 814 TAACCCCGGACCTCGAAACATGAGACCTGAGCTTAAAGATGAGGAGCTATGAGCAATA 755
QY 2571 CAGGCAAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGG 2630
Db 754 CAGGCAAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGG 695
QY 2631 ACACCTGTGGAGAGCTGGAGAGTTAAGAAACACAGAGGTGGACCTCCAAACACATAG 2690
Db 694 WCAAYTGTGGAGAGCTGGAGAGTTAAGAAACACAGAGGTGGACCTCCAAACACATAG 635
QY 2691 AGGCATCACCTGACTGCACAGCAATGAAACCACATGTGGG-TGATTTCCAGC-AGACCT 2748
Db 634 AGGCATCACCTGACTGCACAGCAATGAAACCACATGTGGGTTGATTTCCAGCTAGACT 575
QY 2749 GTGCTATTGCGAGAGGCTCGAGAAAGCAAGACGCACTCTCAGTCAACATGACAGATT 2808
Db 574 KTGTCTTTTGGCAGGAGGCTCGAGAAAGCAAGACGCACTCTCAGTCAACATGACAGATT 515
QY 2809 CTGAGAGATAACACAGAGAGCAGAGATAACTTTCAGGAGTCCATTTTGGCCCTGCTTT 2868
Db 514 CTGAGAGATAACACAGAGAGCAGAGATAACTTTCAGGAGTCCATTTTCCCTCTCTTT 455
QY 2869 TGCTTTGGATTATACCTCACAGCTGCGCAAAATGCAATTTTTCGTATCAAAAAGTCAAC 2928
Db 454 TGCTTTGGATTATACCTCACAGCTGCGCAAAATGCAATTTTTCGTATCAAAAAGTCAAC 395
QY 2929 ACTTACCTTCCCGAGAGCTCACAAGAGAAACGAGAGAGCGAGAGAGAGATTTCC 2988
Db 394 ACTTACCTTCCCGAGAGCTCACAAGAGAAACGAGAGAGCGAGAGAGATTTCC 335
QY 2989 TTGGAATTTCTCCCAAGGGCGAAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAG 3048
Db 334 TTGGAATTTCTCCCAAGGGCGAAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAG 275
QY 3049 TCTGTTCTAAATCTCTTTATTTCTTTTGGTTTGTCAAAA-----GAAGGAACTAAGAA 3102
Db 274 TCTGTTCTAAATCTCTTTATTTCTTTTGGTTTGTCAAAA-----GAAGGAACTAAGAA 215
QY 3103 CGAGGAC-AGAGGCAACCTGGAGAGGCTGAAAACAGTGCAGAGAGCTTTGCAATGAGTC 3161
Db 214 CGAGGAC-AGAGGCAACCTGGAGAGGCTGAAAACAGTGCAGAGAGCTTTTCAATGAGTC 155
QY 3162 AGTAGCAAAAAGAGATGACATTTACCTAGCATTATAAACCTGGTTCCTCTGAAGAAA 3221
Db 154 WGTAGCAAAAAGAGATGACATTTTCCTAGCAATATAAACCTGGTTCCTCTGAAGAAA 95
QY 3222 CTGCTTCTATATATATATATATATATATATATATATATATATATATATATATATATAT 3281
```

```
Db 94 CTGCTTCATTCTATATATGTTACTATTATACATTAATCAACATGGGAGCTTTTGGGCGN 35
QY 3282 ACCTAATAGAAATCCCAATTTTCAGGAGTGGT 3315
Db 34 NCCTAATAGTATCCCAWTTTCWGGAGTGGT 1

BU956086 948 bp mRNA linear EST 21-OCT-2002
AGENCOURT_10621456 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6729576 5', mRNA sequence.
BU956086
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 948)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3052 row: p column: 23
High quality sequence stop: 552.
Location/Qualifiers
1. .948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6729576"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 294 a 201 c 226 g 222 t 5 others
ORIGIN
Query Match 24.2%; Score 812.8; DB 13; Length 948;
Best Local Similarity 97.9%; Pred. No. 1.2e-157;
Matches 853; Conservative 0; Mismatches 15; Indels 3; Gaps 3;
QY 2484 GAGGAGCTGCAAGGTTACAGCAGTGTAAACCCCGACTCGAAACATGGACCTGGACT 2543
Db 1 GAGGAGCTGCAAGGTTACAGCAGTGTAAACCCCGACTCGAAACATGGACCTGGACT 60
QY 2544 TAAAGATGAGGAAGCTATGAGCAATACAGGAGTTCAGCGTCGAAAGTGCCAGAAAT 2603
Db 61 TAAAGATGAGGAAGCTATGAGCAATACAGGAGTTCAGCGTCGAAAGTGCCAGAAAT 120
QY 2604 GAAGAGACCTCTTCCAAATCATCTGGGACAACTGTGGGAGGCTGGGAAGTTAAGAAAC 2663
Db 121 GAAGAGACCTCTTCCAAATCACTGGGACAACTGTGGGAGGCTGGGAAGTTAAGAAAC 180
QY 2664 AACAGAGGTGAGCTCCAAAACATAGAGGCATCACTGACTGCACAGCAATGAAAC 2723
```

```
181 AACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGACTGCACAGCAATGAAAAAC 240
2724 CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGAGAAAGCAGCAC 2783
241 CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGAGAAAGCAGCAC 300
2784 GCACCTCTCAGTCAACATGACAGATTCTGGAGGATAACCCAGCAGGAGCAGAGATAACTCA 2843
301 GCACCTCTCAGTCAACATGACAGATTCTGGAGGATAACCCAGCAGGAGCAGAGATAACTCA 360
2844 GGAAGTCGATTTTGGCCCTGCTTTGCTTTGGATTATACCTCACCAGTCACAAAAATG 2903
361 GGAAGTCGATTTTGGCCCTGCTTTGCTTTGGATTATACCTCACCAGTCACAAAAATG 420
2904 CATTTTTCGTATCAAAAAAGTCAACCACTAACCTCCCTCCAGCAAGCTCAAAAAGGAAACG 2963
421 CATTTTTCGTATCAAAAAAGTCAACCACTAACCTCCCTCCAGCAAGCTCAAAAAGGAAACG 480
2964 GAGAGAGCAGGAGGAGAGATTTCCTTGGAAATTTCTCCAGGCGGAAAGTCATTGGAA 3023
481 GAGAGAGCAGGAGGAGAGATTTCCTTGGAAATTTCTCCAGGCGGAAAGTCATTGGAA 540
3024 TTTTAAATCAT-AGGGGAAAAAGCAGTCTCTGTTCTAAATCTCTTATTCTTTTGGTTGT 3082
541 TTTTAAATCATNAGGGGAAAAAGCAGTCTCTGTTCTAAATCTCTTATTCTTTTGGTTGT 600
3083 CACAAAGAGGAACCTAAGAGCAGGAGCAGGAGCAACGCTGGAGAGCTGAAACAGTGCAG 3142
601 CACAAAGAGGAACCTAAGAGCAGGAGCAGGAGCAACGCTGGAGAGCTGAAACAGTGCAG 660
3143 AGACGTTTGCACATGAGTCAGTAGCAGCAAAAGAGATGACATTTACCTAGCATTATAAAC 3202
661 AGACGTTTGCACATGAGTCAGTAGCAGCAAAAGAGATGACATTTACCTAGCATTATAAAC 720
3203 CTGGTTCCTCTGAAGAACTGCTTATGATATATGATGACTATTATACATGTAATCAA 3262
721 CTGGTTCCTCTGAAGAACTGCTTATGATATATGATGACTATTATACATGTAATCAA 780
3263 CATGG-GAACTTTTGGGGAACCTAATAGAAATCCCAATTTTCAGG-AGTGGTGTGTC 3320
781 CATGGGAACCTTTTANGGGAGCCCTATAGAAATCCCAATTTTCAGGGAAGTGGTGTGTC 840
3321 AATAAAGCTCTGTGGCCAGTGTAAAGAAA 3351
841 AATAAAGCTCTGTGGCCAGTGTAAAGAAA 871

RESULT 13
BQ278788
LOCUS BQ278788
DEFINITION BQ278788 1066 bp mRNA linear EST 07-MAY-2002
5', mRNA sequence.
ACCESSION BQ278788
VERSION BQ278788.1 GI:20488996
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1066)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2044 row: k column: 08
```



Query Match 23.7%; Score 798.6; DB 13; Length 885;

Best Local Similarity 98.1%; Pred. No. 1e-154; Matches 829; Conservative 0; Mismatches 14; Indels 2; Gaps 2;		
QY	2484	GAGGAGCTGCAAGGGTTAACAAGCAGTGTAAACCCCGGACTCGAAACATGTGGACCTGGGACT 2543
DB	1	GAGGAGCTGCAAGGGTTAACAAGCAGTGTAAACCCCGGACTCGAAACATGTGGACCTGGGACT 60
QY	2544	TAAAGATGAGGAAGCTATGAGCAATACAGGAGTTCAGCGTTCGAAAGTGCCCAAGAAAT 2603
DB	61	TAAAGATGAGGAAGCTATGAGCAATACAGGAGTTCAGCGTTCGAAAGTGCCCAAGAAAT 120
QY	2604	GAAGAGACCTTCTTCCAAATCACTGGGCAAACTGTGGGAAGGCTGGGAAGTTTGAAGAAC 2663
DB	121	GAAGAGACCTTCTTCCAAATCACTGGGCAAACTGTGGGAAGGCTGGGAAGTTTGAAGAAC 180
QY	2664	ACAGAGGTGGACCTCCAAAACATAGAGGCATCACTGCTGCTGACAGGCATGAAAGAAC 2723
DB	181	ACAGAGGTGGACCTCCAAAACATAGAGGCATCACTGCTGCTGACAGGCATGAAAGAAC 240
QY	2724	CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGAGGCTTGAGAAAGCAAGCAC 2783
DB	241	CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGAGGCTTGAGAAAGCAAGCAC 300
QY	2784	GCACTCTCAGTCAACATGACAGATTTCTGAGGATTAACAGCAGGAGGACAGATTAATTCA 2843
DB	301	GCACTCTCAGTCAACATGACAGATTTCTGAGGATTAACAGCAGGAGGACAGATTAATTCA 360
QY	2844	GGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTATACCTCACCAGCTGCACAAAATG 2903
DB	361	GGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTATACCTCACCAGCTGCACAAAATG 420
QY	2904	CATTTTTCGTATCAAAAAGTCACTAACCCCTCCCGAGAGCTCAAAAGGAAAACG 2963
DB	421	CATTTTTCGTATCAAAAAGTCACTAACCCCTCCCGAGAGCTCAAAAGGAAAACG 480
QY	2964	GAGAGAGCAGAGGAGAGATTTCTTGGAAATTTCTCCAGAGGCGAAAGTCATTGAA 3023
DB	481	GAGAGAGCAGAGGAGAGATTTCTTGGAAATTTCTCCAGAGGCGAAAGTCATTGAA 540
QY	3024	TTTTTAAATCATAGGGAAGAGCAGTCTGTTCTTAATCCTCTTATTTCTTTTGGTTTCTC 3083
DB	541	TTTTTAAATCATAGGGAAGAGCAGTCTGTTCTTAATCCTCTTATTTCTTTTGGTTTCTC 600
QY	3084	ACAAAGAGAACTAAGACGAGCAGAGGCAACCTGGAGAGGCTGAAAACAGTGCAGA 3143
DB	601	ACAAAGAGAACTAAGACGAGCAGAGGCAACCTGGAGAGGCTGAAAACAGTGCAGA 660
QY	3144	GACGTTTGACATGAGTCAAGTACGACAAAGAGATGACATTTACCTAGCATTATAAACCC 3203
DB	661	GACGTTTGACATGAGTCAAGTACGACAAAGAGATGACATTTACCTAGCATTATAAACCC 720
QY	3204	TGTTGGCTCTGAAGAACTGCTTCAATTTGTATATATGACTATTTACATGTAATCAAC 3263
DB	721	TGTTGGCTCTGAAGAACTGCTTCAATTTGTATATATGACTATTTACATGTAATCAAC 780
QY	3264	ATGG-GAACTTTTAGGGAA-CCTAATAAGAAATCCCAATTTTACAGGAGTGGTGTGCA 3321
DB	781	ATGGGAACTTTTAGGGAACTTATAGAAATCCCAATTTTACAGGAGTGGTGTG 840
QY	3322	ATAAA 3326
DB	841	GTCAA 845
RESULT 15		
BQ884082		
LOCUS		
DEFINITION		BQ884082 904 bp mRNA linear EST 16-AUG-2002
ACCESSION		AGENCOURT_8119366 Lupski dorsal root ganglion Homo sapiens cDNA
VERSION		clone IMAGE:6179939 5', mRNA sequence.
KEYWORDS		BQ884082
SOURCE		BQ884082.1 GI:22276090
		EST.
		Homo sapiens (human)

Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
1 (bases 1 to 904)		
NIH-MGC http://imgc.ncbi.nih.gov/.		
National Institutes of Health, Mammalian Gene Collection (MGC)		
Unpublished		
Contact: Robert Strausberg, Ph.D.		
Email: rsb@nih.gov		
Tissue Procurement: Dr. James R. Lupski		
CDNA Library Preparation: Life Technologies, Inc.		
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)		
Agencourt Bioscience Corporation		
Clone distribution: MGC clone distribution information can be		
found through the I.M.A.G.E. Consortium/LLNL at:		
http://imgc.llnl.gov		
Plate: LLM13561 Row: 0 Column: 12		
High quality sequence stop: 657.		
Location/Qualifiers		
1..904		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/clone="IMAGE:6179939"		
/sex="male"		
/tissue_type="dorsal root ganglia"		
/dev_stage="adult, 36 yr"		
/lab_host="DH10B"		
/clone_lib="Lupski dorsal root ganglion"		
Note: Vector: pCMV-SPORT6 (Life Technologies); Site 1:		
NotI; Site 2: SalI; cDNA made by oligo-dT priming.		
Directionally cloned using the following adaptors:		
5'-TGACACCAAGGCTCCG-3' and		
5'-GATGTTCTAGATCCGACGCGCCCT(15)-3'. Size selected >		
1 kb for average insert length 1.7 kb. This is a primary		
library, non-amplified. Library constructed by Life		
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor		
College of Medicine) and is available through Life		
Technologies."		
BASE COUNT		252 a 265 c 240 g 147 t
ORIGIN		
Query Match		23.5%; Score 791.4; DB 13; Length 904;
Best Local Similarity		96.7%; Pred. No. 3.2e-153;
Matches		862; Conservative 0; Mismatches 21; Indels 8; Gaps 5;
QY	1658	GCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCGTAGGCTGGGTGATGCCGCC 1717
DB	1	GCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCGTAGGCTGGGTGATGCCGCC 60
QY	1718	AGCCCCGAAACCTCACCAGCGGCCTGGCCAGGGCCCTGAGGACCAAGATGACAAGG 1777
DB	61	AGCCCCGAAACCTCACCAGCGGCCTGGCCAGGGCCCTGAGGACCAAGATGACAAGG 120
QY	1778	ATGGTGGGGACTTCAGTGGCCTGAGGGCTTCCGACTACTCAGCGCCCAACCCATTA 1837
DB	121	ATGGTGGGGACTTCAGTGGCCTGAGGGCTTCCGACTACTCAGCGCCCAACCCATTA 180
QY	1838	AAGTGACACATCGGTGTACATCTAGAGAACACACAGTCCAGTGTGACCTGGACCTGT 1897
DB	181	AAGTGACACATCGGTGTACATCTAGAGAACACACAGTCCAGTGTGACCTGGACCTGT 240
QY	1898	ACAAGTCCCTGCGAGGCTTGGAAAGACCAAGCTGCACATCGACCAAGATTTGAACCC 1957
DB	241	ACAAGTCCCTGCGAGGCTTGGAAAGACCAAGCTGCACATCGACCAAGATTTGAACCC 300
QY	1958	TGCGAACAATAATTAAGACCTGAGGAGTCCGAGGTACCTGAGAAAGAGCGGCCAG 2017
DB	301	TGCGAACAATAATTAAGACCTGAGGAGTCCGAGGTACCTGAGAAAGAGCGGCCAG 360
QY	2018	AAGAAATGTGCTGTCAAAAATCAGTACCAACCCAGCACAAGGCGCCCTCAAGCACA 2077
DB	361	AAGAAATGTGCTGTCAAAAATCAGTACCAACCCAGCACAAGGCGCCCTCAAGCACA 420

```
Qy 2078 GAGGCTCCAGTCTGCATCCCTTTTCAGGAAGGGCTTGCAGAGAGGACAAGGTGTGGCTGT 2137
Db |||||
Qy 421 GAGGCTCCAGTCTGCATCCCTTTTCAGGAAGGGCTTGCAGAGAGGACAAGGTGTGGCTGT 480
Db |||||
Qy 2138 TGGGGAGCAGAGCGGAGAGAACTCCGCAAGCTGTCTCAGCGGCTTGCAGAACACG 2197
Db |||||
Qy 481 TGGGGAGCAGAGCGGAGAGAACTCCGCAAGCTGTCTCAGCGGCTTGCAGAACACG 540
Db |||||
Qy 2198 ACAGCTGCAGCATGCCAGGCTTCCAGCTGTCTTCCAGCAGCAACAGCACTGGCAGACGG 2257
Db |||||
Qy 541 ACAGCTGCAGCATGCCAGGCTTCCAGCTGTCTTCCAGCAGCAACAGCACTGGCAGACGG 600
Db |||||
Qy 2258 CGCCTTCTGACACTGGGGCTTCTGTGCTGTGACCCAGCGCCAAACAATACACGTACT 2317
Db |||||
Qy 601 CGCCTTCTGACACTGGGGCTTCTGTGCTGTGACCCAGCGCCAAACAATACACGTACT 660
Db |||||
Qy 2318 GGTGATGAGGACCATCAATCAGACTCACAAATTCCTTCTGTGTAATTTGCCACTGGCT 2377
Db |||||
Qy 661 GGTGATGAGGACCATCAATCAGACTCACAAATTCCTTCTGTGTAATTTGCCACTGGCT 720
Db |||||
Qy 2378 TCCTAGACTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAAGTGAACACAC 2437
Db |||||
Qy 721 TCCTAGACTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAAGTGAACACAC 780
Db |||||
Qy 2438 T-GGACAGGGATGCTCTCAA-CCAGCTACAGT-ACAGCTCATGG--AGCTGAGAGCTG 2492
Db |||||
Qy 781 TGGGACAGGGATGCTCTCAA-CCAGCTACAGTAAACAGCTCATGGAGCTGGGAGAGCTG 840
Db |||||
Qy 2493 CAAGGGTT---ACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGG 2540
Db |||||
Qy 841 CCAGGGTTACAAACCAAGTGTACCCCGGAGCTCCAAACATGGAACTGG 891
Db |||||
```

Search completed: February 14, 2004, 19:38:17  
Job time : 6219 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 22:52:56 , Search time 43 Seconds  
(without alignments)  
856.057 Million cell updates/sec

Title: US-10-025-966A-6  
Perfect score: 4750  
Sequence: 1 MGPSLVCLLSATVFLSLG.....PEMKRPSKSLGQLWEGWG 870

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/FCFUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	4723.5	99.4	867	4	US-09-996-243-84
2	3776	79.5	690	4	US-03-886-319A-36
3	3179.5	66.9	871	4	US-09-773-426A-3
4	3121.5	65.7	867	4	US-09-668-673B-2
5	3063	64.5	1611	4	US-09-668-673B-16
6	2093	44.1	455	4	US-09-668-673B-14
7	1702	35.8	470	4	US-09-886-319A-35
8	1610.5	33.9	360	4	US-09-810-347-4
9	1594.5	33.6	372	4	US-09-810-347-2
10	1584	33.3	709	4	US-09-668-673B-3
11	1494	31.5	307	4	US-09-810-347-5
12	1187	25.0	309	4	US-09-810-347-6
13	827.5	17.4	510	1	US-08-484-493-11
14	827.5	17.4	510	2	US-08-484-494-11
15	827.5	17.4	510	2	US-08-345-212-11
16	827.5	17.4	510	3	US-09-249-003-11
17	827.5	17.4	510	4	US-09-685-844-11
18	565	11.9	160	4	US-09-668-673B-7
19	507	10.7	98	4	US-09-668-673B-22
20	405	8.5	96	4	US-09-668-673B-20
21	399.5	8.4	196	4	US-09-668-673B-18
22	370.5	7.8	520	4	US-09-773-426A-10
23	325	6.8	552	4	US-09-773-426A-9
24	281.5	5.9	406	4	US-09-227-357-187
25	272.5	5.7	556	1	US-08-445-586-2
26	271.5	5.7	554	1	US-08-445-586-7
27	238.5	5.0	583	1	US-08-445-586-9

Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 10, Appli

US-09-646-683-1  
US-08-484-493-2  
US-08-484-494-2  
US-08-345-212-2  
US-09-249-003-2  
US-09-685-844-2  
US-08-484-493-14  
US-08-484-494-14  
US-08-345-212-14  
US-09-249-003-14  
US-09-685-844-14  
US-09-252-991A-25705  
US-08-484-493-15  
US-08-484-494-15  
US-08-345-212-15  
US-09-249-003-15  
US-09-685-844-15  
US-08-445-586-10

ALIGNMENTS

RESULT 1

US-09-996-243-84  
; Sequence 84, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gottard, Audrey E.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavich, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC13  
; CURRENT APPLICATION NUMBER: US/09/996.243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28

1	1	PRIOR APPLICATION NUMBER: 60/084600	1	1	PRIOR FILING DATE: 1998-06-17
2	2	PRIOR FILING DATE: 1998-05-07	2	2	PRIOR APPLICATION NUMBER: 60/089655
3	3	PRIOR APPLICATION NUMBER: 60/087106	3	3	PRIOR FILING DATE: 1998-06-17
4	4	PRIOR FILING DATE: 1998-05-28	4	4	PRIOR APPLICATION NUMBER: 60/089801
5	5	PRIOR APPLICATION NUMBER: 60/087607	5	5	PRIOR FILING DATE: 1998-06-18
6	6	PRIOR FILING DATE: 1998-06-02	6	6	PRIOR APPLICATION NUMBER: 60/089907
7	7	PRIOR APPLICATION NUMBER: 60/087609	7	7	PRIOR FILING DATE: 1998-06-18
8	8	PRIOR FILING DATE: 1998-06-02	8	8	PRIOR APPLICATION NUMBER: 60/089908
9	9	PRIOR APPLICATION NUMBER: 60/087759	9	9	PRIOR FILING DATE: 1998-06-18
10	10	PRIOR FILING DATE: 1998-06-02	10	10	PRIOR APPLICATION NUMBER: 60/089947
11	11	PRIOR APPLICATION NUMBER: 60/087827	11	11	PRIOR FILING DATE: 1998-06-19
12	12	PRIOR FILING DATE: 1998-06-03	12	12	PRIOR APPLICATION NUMBER: 60/089948
13	13	PRIOR APPLICATION NUMBER: 60/088021	13	13	PRIOR FILING DATE: 1998-06-19
14	14	PRIOR FILING DATE: 1998-06-04	14	14	PRIOR APPLICATION NUMBER: 60/089952
15	15	PRIOR APPLICATION NUMBER: 60/088025	15	15	PRIOR FILING DATE: 1998-06-19
16	16	PRIOR FILING DATE: 1998-06-04	16	16	PRIOR APPLICATION NUMBER: 60/090246
17	17	PRIOR APPLICATION NUMBER: 60/088026	17	17	PRIOR FILING DATE: 1998-06-22
18	18	PRIOR FILING DATE: 1998-06-04	18	18	PRIOR APPLICATION NUMBER: 60/090252
19	19	PRIOR APPLICATION NUMBER: 60/088028	19	19	PRIOR FILING DATE: 1998-06-22
20	20	PRIOR FILING DATE: 1998-06-04	20	20	PRIOR APPLICATION NUMBER: 60/090254
21	21	PRIOR APPLICATION NUMBER: 60/088029	21	21	PRIOR FILING DATE: 1998-06-22
22	22	PRIOR FILING DATE: 1998-06-04	22	22	PRIOR APPLICATION NUMBER: 60/090349
23	23	PRIOR APPLICATION NUMBER: 60/088030	23	23	PRIOR FILING DATE: 1998-06-23
24	24	PRIOR FILING DATE: 1998-06-04	24	24	PRIOR APPLICATION NUMBER: 60/090355
25	25	PRIOR APPLICATION NUMBER: 60/088033	25	25	PRIOR FILING DATE: 1998-06-23
26	26	PRIOR FILING DATE: 1998-06-04	26	26	PRIOR APPLICATION NUMBER: 60/090429
27	27	PRIOR APPLICATION NUMBER: 60/088326	27	27	PRIOR FILING DATE: 1998-06-24
28	28	PRIOR FILING DATE: 1998-06-04	28	28	PRIOR APPLICATION NUMBER: 60/090431
29	29	PRIOR APPLICATION NUMBER: 60/088167	29	29	PRIOR FILING DATE: 1998-06-24
30	30	PRIOR FILING DATE: 1998-06-05	30	30	PRIOR APPLICATION NUMBER: 60/090435
31	31	PRIOR APPLICATION NUMBER: 60/088202	31	31	PRIOR FILING DATE: 1998-06-24
32	32	PRIOR FILING DATE: 1998-06-05	32	32	PRIOR APPLICATION NUMBER: 60/090444
33	33	PRIOR APPLICATION NUMBER: 60/088212	33	33	PRIOR FILING DATE: 1998-06-24
34	34	PRIOR FILING DATE: 1998-06-05	34	34	PRIOR APPLICATION NUMBER: 60/090445
35	35	PRIOR APPLICATION NUMBER: 60/088217	35	35	PRIOR FILING DATE: 1998-06-24
36	36	PRIOR FILING DATE: 1998-06-05	36	36	PRIOR APPLICATION NUMBER: 60/090472
37	37	PRIOR APPLICATION NUMBER: 60/088655	37	37	PRIOR FILING DATE: 1998-06-24
38	38	PRIOR FILING DATE: 1998-06-09	38	38	PRIOR APPLICATION NUMBER: 60/090535
39	39	PRIOR APPLICATION NUMBER: 60/088734	39	39	PRIOR FILING DATE: 1998-06-24
40	40	PRIOR FILING DATE: 1998-06-10	40	40	PRIOR APPLICATION NUMBER: 60/090540
41	41	PRIOR APPLICATION NUMBER: 60/088738	41	41	PRIOR FILING DATE: 1998-06-24
42	42	PRIOR FILING DATE: 1998-06-10	42	42	PRIOR APPLICATION NUMBER: 60/090542
43	43	PRIOR APPLICATION NUMBER: 60/088742	43	43	PRIOR FILING DATE: 1998-06-24
44	44	PRIOR FILING DATE: 1998-06-10	44	44	PRIOR APPLICATION NUMBER: 60/090557
45	45	PRIOR APPLICATION NUMBER: 60/088810	45	45	PRIOR FILING DATE: 1998-06-24
46	46	PRIOR FILING DATE: 1998-06-10	46	46	PRIOR APPLICATION NUMBER: 60/090676
47	47	PRIOR APPLICATION NUMBER: 60/088824	47	47	PRIOR FILING DATE: 1998-06-25
48	48	PRIOR FILING DATE: 1998-06-10	48	48	PRIOR APPLICATION NUMBER: 60/090678
49	49	PRIOR APPLICATION NUMBER: 60/088826	49	49	PRIOR FILING DATE: 1998-06-25
50	50	PRIOR FILING DATE: 1998-06-10	50	50	PRIOR APPLICATION NUMBER: 60/090690
51	51	PRIOR APPLICATION NUMBER: 60/088858	51	51	PRIOR FILING DATE: 1998-06-25
52	52	PRIOR FILING DATE: 1998-06-11	52	52	PRIOR APPLICATION NUMBER: 60/090694
53	53	PRIOR APPLICATION NUMBER: 60/088861	53	53	PRIOR FILING DATE: 1998-06-25
54	54	PRIOR FILING DATE: 1998-06-11	54	54	PRIOR APPLICATION NUMBER: 60/090695
55	55	PRIOR APPLICATION NUMBER: 60/088876	55	55	PRIOR FILING DATE: 1998-06-25
56	56	PRIOR FILING DATE: 1998-06-11	56	56	PRIOR APPLICATION NUMBER:

; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 4; Length 867;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 MGPPSLVLCLSATVFSILGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDOVLG 60
DB 1 MGPPSLVLCLSATVFSILGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDOVLG 60
QY 61 MQVMNKRIRIEQGAHFINAFVITPMCCPSRSSILTKGYVHNENTYNNENSSPSWQA 120
DB 61 MQVMNKRIRIEQGAHFINAFVITPMCCPSRSSILTKGYVHNENTYNNENSSPSWQA 120
QY 121 QHESFTAVYLNSTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLKNRFTNYTLCRGVK 180
DB 121 QHESFTAVYLNSTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLKNRFTNYTLCRGVK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSLRFP 240
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSLRFP 240
QY 241 NASCHITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASCHITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELENTYIVVTADHGYHIGFGLVKGSMPEYFDIRVPVYVRGNVEAGCLNPHIV 360
DB 301 LVETGELENTYIVVTADHGYHIGFGLVKGSMPEYFDIRVPVYVRGNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPVNFHLKXKWRVWDSFLVERKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPVNFHLKXKWRVWDSFLVERKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKOLCQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
DB 421 HKRDNKVDQAQENFLPKYQVRKOLCQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
QY 481 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
DB 481 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
QY 541 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 660
DB 541 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 660
QY 601 HRCYILENTVQCDLDLYKSLQAWKHKLHIDHEITLQNKIKNLRVGRHLKXKRPBEC 720
DB 601 HRCYILENTVQCDLDLYKSLQAWKHKLHIDHEITLQNKIKNLRVGRHLKXKRPBEC 720
QY 661 DCHKISYHTQHKGRKLRHSGSLHPRFKGLQEKDKVLLREQKRRKLLKRLQNNNTC 840
DB 661 DCHKISYHTQHKGRKLRHSGSLHPRFKGLQEKDKVLLREQKRRKLLKRLQNNNTC 840
QY 721 SMPGLTCTHNDQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLFCEPATGFLE 870
DB 721 SMPGLTCTHNDQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLFCEPATGFLE 870
QY 841 QYRQORRWKPEMKGPSSKSLGQWEGWEG 870
DB 841 QYRQORRWKPEMKGPSSKSLGQWEGWEG 870
QY 898 QYRQORRWKPEMKGPSSKSLGQWEGWEG 867
DB 898 QYRQORRWKPEMKGPSSKSLGQWEGWEG 867
```

RESULT 2

US-09-886-319A-36  
; Sequence 36, Application US/09886319A  
; Patent No. 6586185  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Eckard  
; APPLICANT: Werner, Sabine  
; APPLICANT: Halle, Jorn-Peter  
; APPLICANT: Regenbogen, Johannes  
; APPLICANT: Goppelt, Andreas  
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
; the Diagnosis or Treatment of Skin Disorders and Wound  
; Healing and for the Identification of Pharmacologically  
; Active Substances  
; FILE REFERENCE: 50125/014002  
; CURRENT APPLICATION NUMBER: US/09/886,319A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 60/222,081  
; PRIOR FILING DATE: 2000-08-01  
; PRIOR APPLICATION NUMBER: DE 10030149.5  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-886-319A-36

Query Match 79.5%; Score 3776; DB 4; Length 690;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSLRFP 240
DB 1 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSLRFP 60
QY 241 NASCHITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 61 NASCHITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 120
QY 301 LVETGELENTYIVVTADHGYHIGFGLVKGSMPEYFDIRVPVYVRGNVEAGCLNPHIV 360
DB 121 LVETGELENTYIVVTADHGYHIGFGLVKGSMPEYFDIRVPVYVRGNVEAGCLNPHIV 180
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPVNFHLKXKWRVWDSFLVERKLL 420
DB 181 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPVNFHLKXKWRVWDSFLVERKLL 240
QY 421 HKRDNKVDQAQENFLPKYQVRKOLCQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
DB 241 HKRDNKVDQAQENFLPKYQVRKOLCQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 300
QY 481 MELGGSRALSNLVPKYQGSGEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
DB 301 MELGGSRALSNLVPKYQGSGEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 360
QY 541 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
DB 361 ALEVDRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 420
QY 601 HRCYILENTVQCDLDLYKSLQAWKHKLHIDHEITLQNKIKNLRVGRHLKXKRPBEC 660
DB 421 HRCYILENTVQCDLDLYKSLQAWKHKLHIDHEITLQNKIKNLRVGRHLKXKRPBEC 480
QY 661 DCHKISYHTQHKGRKLRHSGSLHPRFKGLQEKDKVLLREQKRRKLLKRLQNNNTC 720
DB 481 DCHKISYHTQHKGRKLRHSGSLHPRFKGLQEKDKVLLREQKRRKLLKRLQNNNTC 540
QY 721 SMPGLTCTHNDQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLFCEPATGFLE 780
DB 541 SMPGLTCTHNDQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLFCEPATGFLE 600
QY 781 YFDLNTDPYQLMNAVNTLDRVLAQLHVLQWELRSCKYKQCNPTRNMDJGLKDGGSYE 840
```



Db 601 YPDLNTPVQLNNAVNTLDRDVLNQLVQLMEARSKGYKQCNPYTRNDWDLGLKGGSYE 660  
QY 841 QYRQFORRWKPEKRPSSKSLQWEGWEG 870  
Db 661 QYRQFORRWKPEKRPSSKSLQWEGWEG 690  
RESULT 3  
US-09-773-426A-3  
; Sequence 3, Application US/09773426A  
; Patent No. 6534302  
; GENERAL INFORMATION:  
; APPLICANT: Gluckman, Maria Alexandra  
; APPLICANT: Williamson, Mark  
; APPLICANT: Teia, Fong-Ying  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el  
; TITLE OF INVENTION: Human Sulfatases (A CIP Application)  
; FILE REFERENCE: 35800/208398(5800-79  
; CURRENT APPLICATION NUMBER: US/09/773,426A  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/495,823  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 871  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-773-426A-3  
Query Match 66.9%; Score 3179.5; DB 4; Length 871;  
Best Local Similarity 65.0%; Pred. No. 9e-302;  
Matches 579; Conservative 120; Mismatches 141; Indels 51; Gaps 12;  
QY 5 SLVLCILSATVPSLGGSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVAVLGT---ELLSGLSTVRSFRGRIQERKNIRPNILVLTDDQDVELGSLQVM 63  
QY 65 NKTRIMEOGGAHFINAFVTPMCCPSRSISLTGKYVHNHNTYNNENCSFSPWQAQHS 124  
Db 64 NKTRIMEOGGAHFINAFVTPMCCPSRSISLTGKYVHNHNTYNNENCSFSPWQAHEP 123  
QY 125 RTFAVLNNGTGYTAFPGKYLYNEYSYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184  
Db 124 RTFAVLNNGTGYTAFPGKYLYNEYSYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183  
QY 185 SDYSKDYLTDLTNDVSFFRTSKQMPYRHPVLMVISHAAPHGPDSPAPQYSLFPNASQ 244  
Db 184 FDYAKDYFTDLITNESINYPFGSKRMYPHPVMMVISHAAPHGPDSPAPQYSLFPNASQ 243  
QY 245 HITPSYNYAPNDKHMIMYTGPMKDIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNDKHMIMYTGPMKDIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
QY 305 GELDNTYIVVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 364  
Db 304 GELENYIIVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 363  
QY 365 LAPTLIDAGLIDPAMWGSILKLDTRPVRNPHLKKGRWRDSDFLVERGKLLHXR 424  
Db 364 LAPTLIDAGLIDPAMWGSILKLDTRPVRNPHLKKGRWRDSDFLVERGKLLHXR 423  
QY 425 NDKVDAEENFLPKYORVKDLCORAEVQACELGOKWOCVEDATGKLLHKCKGPMRLG 484  
Db 424 ESSKNTQSNHLPKYERVELKCOQARYQACELGOKWOCIEDTSGKLLHCKGPMRLG 483  
QY 485 GSR-ALSNLVPKYGGSEACTCDSDGYKLSLAGRR-KKLFXK-----KYKASYVRSRI 537  
Db 484 TVRQSTNLYARGFHDKDKCSKESYGRASRQKRSQFLRNQGTPKYKPRFVHTROT 543  
QY 536 RSVAIEDVGRVYHVLGD-----AAQPNLTKEHWP--APEDQDDKGG-----DFS 583

Db 544 RSLSVFEFEIYDINLEEEELQVLPRIAKRHBGKPRDLOASSGNGRMLADSS 603  
QY 584 GTGGLPDYSAANDIKVTHRCYILENDTVQDDLDLYKSLQAWKDKHLHIDHETIQNKIK 643  
Db 604 NAVGPP-----TTVRVTHKCFILPNDSIHCERELYQSARAWDKHAYIDKEIEALQDKIK 658  
QY 644 NLREVRGHLKKKPEECDCIKSYHTQHKGRLLKHR--GSSLHPFRKGLQEKD-KVWLLRE 700  
Db 659 NLREVRGHLKKRKPPEECSCSKQSYNKEGVKKQEKLSHLHPFKEAAQEVDSKLOLFXE 718  
QY 701 -QKRKKKLLKRLQLQNDTCSMPGLTCTHNDQNHQWQTAPFTWLTGPFCACTGANNNTYWC 759  
Db 719 NNRKKRKEKEKQRQKGECSLPGLTCTHNDNHWQTAPFTWLTGPFCACTGANNNTYWC 778  
QY 760 MRTINETHNPLFCFEPATGFLFVFDLNTDPVQLMNAVNTLDRVLNOLHVQLMELRSCKGY 819  
Db 779 LRIVNETHNPLFCFEPATGFLFVFDLNTDPVQLMNAVNTLDRVLNOLHVQLMELRSCKGY 838  
QY 820 KQCNPRTRNMDLGLKGGSYEQYRQFRKWPMPKRPSSKSLGQWEGWEG 870  
Db 839 KQCNPRKQNLVGNKGGSYDLHR-----GQLWDGWEG 871  
RESULT 4  
US-09-668-673B-2  
; Sequence 2, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 867  
; TYPE: PRT  
; ORGANISM: Quails  
US-09-668-673B-2  
Query Match 65.7%; Score 3121.5; DB 4; Length 867;  
Best Local Similarity 64.1%; Pred. No. 4.2e-296;  
Matches 572; Conservative 117; Mismatches 145; Indels 59; Gaps 10;  
QY 5 SLVLCILSATVPSLGGSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVAVLGT---ELLSHSTLTKSLRFRGVQQRKNIRPNILVLTDDQDVELGSLQVM 63  
QY 65 NKTRIMEOGGAHFINAFVTPMCCPSRSISLTGKYVHNHNTYNNENCSFSPWQAQHS 124  
Db 64 NKTRIMEOGGAHFINAFVTPMCCPSRSISLTGKYVHNHNTYNNENCSFSPWQAHEP 123  
QY 125 RTFAVLNNGTGYTAFPGKYLYNEYSYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184  
Db 124 RTFAVLNNGTGYTAFPGKYLYNEYSYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183  
QY 185 SDYSKDYLTDLTNDVSFFRTSKQMPYRHPVLMVISHAAPHGPDSPAPQYSLFPNASQ 244  
Db 184 FDYAKDYFTDLITNESINYPFGSKRMYPHPVMMVISHAAPHGPDSPAPQYSLFPNASQ 243  
QY 245 HITPSYNYAPNDKHMIMYTGPMKDIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNDKHMIMYTGPMKDIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
QY 305 GELDNTYIVVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 364

Db 304 GELENTYIYIADHGYHIGQGLVKGKSNPYDFDTRVFFIRGSPVSGSVVPOIVLNID 363  
QY 365 LAPPTILDIAGLDPADMDGKSLKLLDTERPVNRFLKKGQVVRDPSPLVERGKLLAKRD 424  
Db 366 LAPPTILDIAGLDPADMDGKSLKLLDTERPVNRFLKKGQVVRDPSPLVERGKLLAKRD 423  
QY 425 NDKVDAQEENFLPKYQKDLQCRAYQACEOLGQKQCVEDATGKLLKHKCKGPRWLG 484  
Db 424 EANKNTQOSNQJPKYVERVELCCQARYQACEOLGQKQCVEDATGKLLKHKCKGPRWLG 483  
QY 485 GSRALSNLV-PRYQGGSEACTCDSDGYKLSLAGRR-----KLFKKKYKASVVRGSI 537  
Db 484 AIRKTRSHSRGYSKDKCNCGDTDFNSPTQKQORQFLRNPQAKYXPRFVHTROT 543  
QY 538 RSVALTEVDORVTHVGLGD-----AAPRNLTKRH-----WGPAPEDQDDKGGGDFS 583  
Db 544 RSLSEFEGEIYDINLESEELQVLKTRSIITKRNAENDKKAETDGPAGDTWADGTDVI 603  
QY 584 GTGGLPDYSAANPIKVTYHCYILENDTVQCDLKYSLQAMKDKHLDHDIETLQNKIK 643  
Db 604 G-----QPSVAVTHKCFILPNDTIRCELYQSARAWDKHAYIDKEIEALQDKIK 655  
QY 644 NLREVRGHLKRRPBCDCHKISYHTQHG-----RLKRGSSLSHPFRKGLQEKD-KVWL 697  
Db 656 NLREVRGHLKRRPBCDCHKISYHTQHG-----RLKRGSSLSHPFRKGLQEKD-KVWL 712  
QY 698 LREOKSKKLRKLLQNLQNDTCSVPLTCTHDNQHWQATPFWTLGPFCACTSANNTY 757  
Db 713 FXENRRKRRKRGKRRKQKGGDCSLPGUTCTFHDNNHWQATPFWTLGPFCACTSANNTY 772  
QY 758 WCMRTINETHNPLFCFATGFLFEDMTDVPQLMNAVNTLDRVNLQHLVQLMELRSCQ 817  
Db 773 WCLRTVNDTHNPLFCFATGFLFEDMTDVPQLMNAVNTLDRVNLQHLVQLMELRSCQ 832  
QY 818 GYKQCNPRTRMDLGLKGGSEYQVQRKRWPMKPKSPSSKSGOLWEGHEG 870  
Db 833 GYKQCNPRTRMDLGLKGGSEYQVQRKRWPMKPKSPSSKSGOLWEGHEG 867

RESULT 5  
US-09-668-673B-16  
Sequence 16, Application US/09668673B  
GENERAL INFORMATION:  
APPLICANT: Emerson, Charles P  
APPLICANT: Dhoot, Gurtej K  
TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
SULFATASES AND FUNCTIONAL EMERSONIC TECHNIQUES FOR  
CHARACTERIZATION OF SUCH PROTEINS  
FILE REFERENCE: PENN-0733  
CURRENT APPLICATION NUMBER: US/09/668,673B  
CURRENT FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,738  
PRIOR FILING DATE: 1998-09-23  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 16  
LENGTH: 1611  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (819)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (840)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (844)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (852)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (858)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (865)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (875)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (878)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (881)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (888)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (896)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (907)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (910)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (915)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (927)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (943)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (945)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (948)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (954)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (959)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (971)  
OTHER INFORMATION: any amino acid  
FEATURE:  
NAME/KEY: UNSURE

;	LOCATION: (974)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1018)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1046)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1080)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1089)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1102)..(1103)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1105)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1121)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1127)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1191)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1199)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1223)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1235)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1250)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1307)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1321)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1356)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1362)	
;	OTHER INFORMATION: any amino acid	
;	FEATURE:	
;	NAME/KEY: UNSURE	
;	LOCATION: (1382)..(1383)	

Db 656 VDSKQLQFKENRRARKERKQRKGECSLGLCTFTHDNNHQTAPFWNLGSFCAC 715  
QY 750 TSANNNTYWCMTINETHNLFCEFATGFLFYDLNTPYQLMNAVNTLDRDVLNQLHVQ 809  
Db 716 TSSNNNTYWCMTINETHNLFCEFATGFLFYDLNTPYQLMNAVNTLDRDVLNQLHVQ 775  
QY 810 LMELRSCKGYKQCNPRTRNMOLGLKDGSSYEQYQFORRWKPEKMRPSSKSLQOLWEGWE 869  
Db 776 LMELRSCKGYKQCNPRTRNMOLGLKDGSSYEQYQFORRWKPEKMRPSSKSLQOLWEGWE 817  
QY 870 G 870  
Db 818 G 818

RESULT 6  
US-09-668-673B-14  
; Sequence 14, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; CURRENT FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; NAME/KEY: UNSURE  
; LOCATION: (142-143)  
; OTHER INFORMATION: any amino acid  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (146)  
; OTHER INFORMATION: any amino acid  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (445)  
; OTHER INFORMATION: any amino acid  
US-09-668-673B-14

Query Match 44.1%; Score 2093; DB 4; Length 455;  
Best Local Similarity 86.1%; Pred. No. 6.6e-196;  
Matches 384; Conservative 16; Mismatches 38; Indels 8; Gaps 4;

QY 431 QENFLPKYQKVLQCORAEYQTAQCEQLGQKQWQVEDATGKLHKCKGPMRL---GGSR 487  
Db 1 QENFLPKYQKVLQCORAEYQTAQCEQLGQKQWQVEDATGKLHKCKGPMRLFGGGGSR 60  
QY 488 ALSNLVPRKYVGGSEACTCDS---GDYKLSIAGRRKKLFKKYKASVRSRSIRSAIEV 544  
Db 61 ALSNLVPRKYVGGSEACTCDSGGGGYKGLAGRR-KLFFKKYKTSYARNRSIRSAIEV 119  
QY 545 DGRVYHVLGDAAPRNLTKRHWPCAPEDQDDXGDFSGTGLPDYSAANPIKVTHRCY 604  
Db 120 DGEIYHVLGDTVPQPNLSKPHXGAXEDQDDXGDFSGTGLPDYSAANPIKVTHRCY 179  
QY 605 ILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLEVRGHLKKRPEECDCBK 664  
Db 180 ILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLEVRGHLKKRPEECDCBK 239  
QY 665 ISYTHQKGRGLKRGSSSLHPRFKGLQEKDKVLLREOKRKKLKLKRLQNNDTCSPMG 724  
Db 240 ISYTHQKGRGLKRGSSSLHPRFKGLQEKDKVLLRDNRNARENCATAQTAAEQR-YVQHAG 298

QY 725 LTCFTHDNOHWQTAPFWNLGPFCACTSANNTYWCMTINETHNLFCEFATGFLFYEDL 784  
Db 299 LTCFTHDNOHWQTAPFWNLGPFCACTSANNTYWCMTINETHNLFCEFATGFLFYEDL 358  
QY 785 NTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMOLGLKDGSSYEQYRQ 844  
Db 359 STDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMOLGLKDGSSYEQYRQ 418  
QY 845 FORRWKPEKMRPSSKSLQOLWEGWEG 870  
Db 419 FORRWKPEKMRPSSKSLQOLWEGWEG 444

RESULT 7  
US-09-886-319A-35  
; Sequence 35, Application US/09886319A  
; Patent No. 6586185  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Eckard  
; APPLICANT: Werner, Sabine  
; APPLICANT: Halle, Jörn-Peter  
; APPLICANT: Regenbogen, Johannes  
; APPLICANT: Goppelt, Andreas  
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
; FILE REFERENCE: 50125/014002  
; CURRENT APPLICATION NUMBER: US/09/886,319A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 60/222,081  
; PRIOR FILING DATE: 2000-08-01  
; PRIOR APPLICATION NUMBER: DE 10030149.5  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; NAME/KEY: VARIANT  
; LOCATION: 371  
; OTHER INFORMATION: Xaa= any amino acid  
US-09-886-319A-35

Query Match 35.8%; Score 1702; DB 4; Length 470;  
Best Local Similarity 75.3%; Pred. No. 1.3e-157;  
Matches 327; Conservative 21; Mismatches 36; Indels 50; Gaps 7;

QY 300 MLVETGELDNITVIVTADGHYHIGQGLVKGKSMPEYEDIRVPFVVRGPNVEAGCLNPHI 359  
Db 1 MLVETGELDNITVIVTADGHYHIGQGLVKGKSMPEYEDIRVPFVVRGPNVEAGCLNPHI 60  
QY 360 VLNIDLAP-----TILDIAGLDIPADMKGKSLKLLDTERPVN 397  
Db 61 VLNIDLGPHTGYRTGHPCHRGREVSQTTGLRAASEPVPLEKEAEG-----GETP-- 113  
QY 398 RFLHKKKQWVRWDSFLVERGKLLHKRDNKDYDAQENFLPKYQKVLQCORAEYQTAQCEQ 457  
Db 114 -----SWREA-----NLLHKREGDKYNAQENFLPKYQKVLQCORAEYQTAQCEQ 159  
QY 458 LGQKQWQVEDATGKLHKCKGPMRL---GGSRALSILVPRKYVGGSEACTCDS---GDY 511  
Db 160 LGQKQWQVEDATGKLHKCKGPMRFGGGGGSRALSILVPRKYVGGSEACTCDSGGGSDY 219  
QY 512 KLSLAGRRKKLPKKYKASVRSRSIRSAIEVDGRVTHVGLGDAAPRNLTKRHWPCAP 571  
Db 220 KLGLAGRR-KLFFKKYKTSYARNRSIRSAIEVDGEIYHVLGDTVPQPNLSKPHWPCAP 278  
QY 572 EPQDDKDGDFSGTGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKHLDH 631

Db 279 EQDDKDGSGFSGTGLPDYSAPNPIKVTTHRCYILNDVTQCDLDLYLQAWDKHLI 338  
Qy 632 DHEIETLQNKILREVGHLLKKEPECDCHKISYHTQHKRGLKRGSSILHPRKGLQE 691  
Db 339 DHEIETLQNKILREVGHLLKKEPECDCHKISYHTQHKRGLKRGSSILHPRKGLQE 399  
Qy 692 KDKVLLREQRKK 705  
Db 399 KDKVLLRDNRARR 412  
RESULT 8  
US-09-810-347-4  
; Sequence 4, Application US/09810347  
; Patent No. 6461847  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001169  
; CURRENT APPLICATION NUMBER: US/09/810,347  
; CURRENT FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Human  
US-09-810-347-4

Query Match 33.9%; Score 1610.5; DB 4; Length 360;  
Best Local Similarity 80.1%; Pred. No. 7.4e-149;  
Matches 286; Conservative 40; Mismatches 28; Indels 3; Gaps 1;  
Qy 5 SLVCLLSATVPSLLGGSAFLSHRLKGRFDRNRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVAVLGT---ELLGSLCSTVRSFRFGRIQOERKNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEGGGAHFINAFVTTMCCPSSRSSIITGKYVHNHNTYNNENCSFSSWQAQHE 124  
Db 64 NKTRKIMEGGGATFINAFVTTMCCPSSRSSMLTGKYVHNHNTYNNENCSFSSWQAHEP 123  
Qy 125 RTPAVYLNSTGYRTAFPGKLYNEVNGSYVPPGKRWGLLKNRFPYNTLCRNGVKEKHG 184  
Db 124 RTPAVYLNSTGYRTAFPGKLYNEVNGSYVPPGKRWGLLKNRFPYNTLCRNGVKEKHG 183  
Qy 185 SDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAAPHGPDSDAPQYSRLFPNNA 244  
Db 184 FDYAKDYFTDLITNESINFYFKSKMYPHRPVLMVISHAAPHGPDSDAPQFSKLYPNASQ 243  
Qy 245 HITPSYNYAPNPKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNPKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTADGHYHIGQFGLVKGKSMYPYDFDIRVPPYVRGPNVEAGCLNPHI 361  
Db 304 GELENTYIIVYTADGHYHIGQFGLVKGKSMYPYDFDIRVPPYVRGPNVEAGCLNPHI 360

RESULT 9  
US-09-810-347-2  
; Sequence 2, Application US/09810347  
; Patent No. 6461847  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001169  
; CURRENT APPLICATION NUMBER: US/09/810,347  
; CURRENT FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Human  
US-09-810-347-2  
Query Match 33.6%; Score 1594.5; DB 4; Length 372;  
Best Local Similarity 80.3%; Pred. No. 2.9e-147;  
Matches 282; Conservative 40; Mismatches 26; Indels 3; Gaps 1;  
Qy 5 SLVCLLSATVPSLLGGSAFLSHRLKGRFDRNRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVAVLGT---ELLGSLCSTVRSFRFGRIQOERKNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEGGGAHFINAFVTTMCCPSSRSSIITGKYVHNHNTYNNENCSFSSWQAQHE 124  
Db 64 NKTRKIMEGGGATFINAFVTTMCCPSSRSSMLTGKYVHNHNTYNNENCSFSSWQAHEP 123  
Qy 125 RTPAVYLNSTGYRTAFPGKLYNEVNGSYVPPGKRWGLLKNRFPYNTLCRNGVKEKHG 184  
Db 124 RTPAVYLNSTGYRTAFPGKLYNEVNGSYVPPGKRWGLLKNRFPYNTLCRNGVKEKHG 183  
Qy 185 SDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAAPHGPDSDAPQYSRLFPNNA 244  
Db 184 FDYAKDYFTDLITNESINFYFKSKMYPHRPVLMVISHAAPHGPDSDAPQFSKLYPNASQ 243  
Qy 245 HITPSYNYAPNPKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNPKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTADGHYHIGQFGLVKGKSMYPYDFDIRVPPYVRGPNVEAGCL 355  
Db 304 GELENTYIIVYTADGHYHIGQFGLVKGKSMYPYDFDIRVPPYVRGPNVEAGCL 354

RESULT 10  
US-09-668-673B-3  
; Sequence 3, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; CURRENT FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 709  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-668-673B-3

Query Match 33.3%; Score 1584; DB 4; Length 709;  
Best Local Similarity 40.1%; Pred. No. 8.9e-146;  
Matches 316; Conservative 118; Mismatches 197; Indels 158; Gaps 17;  
Qy 45 NIILVLTDDQDVELGSMQVMNKTRIMEGGGAHFINAFVTTMCCPSSRSSIITGKYVHNH 104  
Db 36 NVILVLTDDQDIELGSMDFMPKTSQIMKRGTEFTSGYVTTTICCPSSRTILTGLVYHNH 95  
Qy 105 NTVNNENCSFSSWQAQHESTRTPAVYLNSTGYRTAFPGKLYNEVNGSYVPPGKRWGLL 164  
Db 96 HVHTNNQNTCTGVEMRKVHEKKSIGVYQLQAGYATYGLYNEIDGSIYPPGDEHAI 155  
Qy 165 KNSRFNYNTLCRNGVKEKHGSDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAA 224





US-08-484-493-11  
; Sequence 11, Application US/08484493  
; Patent No. 5728381  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; TITLE OF INVENTION: IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,493  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8416Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 510 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-484-493-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;  
Best Local Similarity 27.5%; Pred. No. 5.6e-72;  
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RPNILVLDQDVELGSMQVKNTRIMEQGAFFINAVFTPMCCPSRSILTKGVH 102  
DB 3 RPNVLLTDDQDELVLGHTPLKTKALICGMGTFFSSAYPSALCCPSRSILTKRYPH 62  
QY 103 NHTYTN--NENGSSPSWAQHSRRTFAVVLNS-TGYR--TAPFGKYLNEYNG-----S 151  
DB 63 NHHVAVNTLEGNCSSKSWQKIQEPNTPAILRSMQGYQTFPPAGKYLNEYGAPDAGGLE 122  
QY 152 YVPPGKWEVGLLKNRFFNYTLCRNGVEKGGSDYSKDYLTDLITNDSVSP--FRSKK 209  
DB 123 HVLPGWSYVALEKNGKYNITLTSINGKARKGENYSVDYLDVLAVNSLDFDYKSNEE 182  
QY 210 MYHRPVLVISHAAPHGEDSPAPQSRLLFPNASQHI TPSVYVAPNPDGHWIMRYT-GPM 268  
DB 183 -----PFFMWIATPAHSPFWTAPQYQKAFQNVFAPRNKFN-INGTNKHWIIRQAKTPM 236  
QY 269 KPIMHFTNNLQRIQLTMSVDDSMETIYNMLVETGELDNTYIVYTADHGYHIQGFGLV 328  
DB 237 TNSSIQFLDNAFRKRQTLTLLSVDDLVEKLVRLEFTGELNNTYIFYTSNGYHTGQFSIP 296

QY 329 KGKSMPEYFEDIRVPFYVGRPNVEAGCLNPHIVLNIDLAPTILDIAGLDI-PADMDGKSIL 387  
DB 297 IDKRQLYFEDIKVPILLVRGPGIKPNQTSKMLVANIDLGP-ILDIAGYDLNKTQMDGMSLL 355  
QY 388 KLLDTERPVNRFHLKKOMRVWRDSFLVERGKLLHKRDNKDVAQEENFLPKYQRVKDLQ 447  
DB 356 PIL---RGASNL-----TWRSDDLVE-----YQEGG----- 373  
QY 448 RAEYQTACEQLGQKQWQVEDATGKLKHKCKGPMRLGGSRALSNLVPKYVGQSEACTCD 507  
DB 374 -----YQEGG----- 378  
QY 508 SGDYKLSLAGRRKGLFKKKYKASYVRSIRSVALEVDGRVTHVGLGDAQAQPNLTKRHW 567  
DB 379 -----RNVY----- 382  
QY 568 PGAPEDQDDKDGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAKDH 627  
DB 383 -----PGVSQCF-----PD 399  
QY 747 CACTSANNNTYCMRTINETHNLFCEP--ATGFLFYEDLNTDPYQLMNAVNTLDRVLN 804  
DB 400 CVCEDAYNNTYACVRTMSALWNLQYCEFPDQBFVEVYVNLTAADPQITNIAKTIDPELIG 459  
QY 805 QLVHQLMELRSCKG 818  
DB 460 KMRYLMMQLQSCSG 473

RESULT 14  
; Sequence 11, Application US/08484494  
; Patent No. 5798239  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; TITLE OF INVENTION: IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,494  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:

```
/ NAME: DiGiglio, Frank S
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 8416Z
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 516-742-4343
/ TELEFAX: 516-742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 510 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-484-494-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 5.6e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RNNILVLTDDQDVBLGSMQVANKTRRIMEQGAHFINAFVTPMCCPSRSILTKYVH 102
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 RNNVLLTDDQDVLGGMTPLKTKALIGEMGTFSSAYVPSALCCPSRASILTGYPH 62
QY 103 NNTYTN--NENCSSPSWQAQHSRTFAVLNS--TGYR--TAFPGKYLNEYNG-----S 151
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 NHHVNVNTEGNCSSKSWKIOEPNTFFAILRSMQGYOTFFFAGKYLNEYCAPDAGGLE 122
QY 152 YPPGKWEVGLLKNSRFNYTLCRNGVKEKGSYKDYLTDLTNDVSVF--FRTSKK 209
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 HVPGLGWSYVALEKNSKYNNTLSINGKARKHGENYSVDYLTDLVANSFLDLYKSNEE 182
QY 210 MYPHRPVLVIVISHAAPHGEDSAPQYSRLFPNASQHTPSVNYAPNPKHWMYRTPM 268
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 -----PFFMMIATPAPHGFWTAPOYQAFQVFAFRNKNFN--IHGTNKGWLIQAQTPM 236
QY 269 KPIHMEFTNMLQRKQLTMSVDDSMETIYNNMLVETGBELDNTIYVYTAHGYHIGQGLV 328
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 TNSSIQFLDNAFRKRWQTLTSSVDLVEKLVKLEFTGELNNTIYFYTSDNGYHTGQSLP 296
QY 329 KGSMPYEDIRVPYVRGPNVEACLPNIVNLIDLAFTILDIAGLDI--PADWDGKSIL 387
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 IDKROLQYEDIKVPLVVRGPGIKPNQTSKMLVANIDLGP--ILDIAGYDLNKTQMDGSL 355
QY 388 KLLDTERPVNRPELKKMKRVRWDSFLVERGKLLHCRDNDKVDQAEENFLPYQYRVDLCQ 447
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 PIL---RGASNL-----THRSVDLVE----- 373
QY 448 RAEYOTACEQLGQKQWQVEDATGKLGKHKRGMELGGSRLSNLPKYQOGSEACTCD 507
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
374 -----YQEG----- 378
QY 508 SGDYKLSLAGRKKLFKKYKASYRSRSIRSVAIEVDGRVYHVLGDAAQPNLTKRHW 567
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 -----RNV----- 382
QY 568 PGAPEDQDDKGGDFSGTGGLPDYGAANPIKVTHRCYILENTVQCOLDLYKSLQAWKH 627
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 ----- 382
QY 628 KLIHDIETLQNKIKNLREVRGHLKKRPEBCDCHKI SYTHQKGRUKRKGSSLHPRK 687
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 -----DPTCPSL----- 390
QY 688 GLQEKDKVLLREOKRKKLKLRLQNNDTCSMPGLT-CFTDNDQHWQAPFWTLGPF 746
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
391 -----PGVSQCF----- 399
QY 747 CACTGANNNTYCMRTINETHNLFCEP--ATGFLYFDLNTDPQLMNVNLTLDRLVNL 804
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 CVCEDAYNTYACVRTMSALMQLQCEFDQDQEVFEVYVNLTAQPDQITNIAKTIDPELLG 459
QY 805 QLHVQLMELRSCCK 818
```

```
Db 460 KXVRLMMLQSCSG 473
::: |||: ||| |
RESULT 15
US-08-345-212-11
; Sequence 11, Application US/08345212
; Patent No. 5932211
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,212
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-345-212-11
```

```
Query Match 17.4%; Score 827.5; DB 2; Length 510;
Best Local Similarity 27.5%; Pred. No. 5.6e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RNNILVLTDDQDVBLGSMQVANKTRRIMEQGAHFINAFVTPMCCPSRSILTKYVH 102
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 RNNVLLTDDQDVLGGMTPLKTKALIGEMGTFSSAYVPSALCCPSRASILTGYPH 62
QY 103 NNTYTN--NENCSSPSWQAQHSRTFAVLNS--TGYR--TAFPGKYLNEYNG-----S 151
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 NHHVNVNTEGNCSSKSWKIOEPNTFFAILRSMQGYOTFFFAGKYLNEYCAPDAGGLE 122
QY 152 YPPGKWEVGLLKNSRFNYTLCRNGVKEKGSYKDYLTDLTNDVSVF--FRTSKK 209
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 HVPGLGWSYVALEKNSKYNNTLSINGKARKHGENYSVDYLTDLVANSFLDLYKSNEE 182
QY 210 MYPHRPVLVIVISHAAPHGEDSAPQYSRLFPNASQHTPSVNYAPNPKHWMYRTPM 268
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 183 -----PFEMMIATPAPHPWTAAPQOKAFQNFVAFPRNKFN- IHGTNKHILIRQAKTEM 236  
QY 269 KPIHMETNMLQKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYTADHGYHIGOPGLV 328  
Db 237 TNSSIQFLDNFRKRWTLLSVDDLVKLVKRLFTGELNNTYIFYTSDNGYHTGQFSLP 296  
QY 329 KGKSMPEYFDIRVPPYVRGPNVEAGCLNPHIVLINIDLAPTILDIAGLDI- PADMDGKSIL 387  
Db 297 IDKROLYEFDIKVELLVGPGIKPNQTSKMLVANIDLGP- ILDIAGYDLNKTQMDGMSLL 355  
QY 388 KLLDTERPVRNFKKQORVWRDSFLVERGKLLHKDNDKVDQAEENFLPKYQRYKDLQ 447  
Db 356 PIL---RGASNL-----TWRSDDLVE----- 373  
QY 448 RAEYQTACEQLGQKWQCVEDATGKLLHKCKGPMRLGGSRAISNLVPKYGGGSEACTCD 507  
Db 374 -----YQEG----- 378  
QY 508 SGDYKLSLAGRRKKLFKKYKASVRSRSIRSVAIEVDGRVHVGLGDAAPRNLTGRHW 567  
Db 379 -----RNV----- 382  
QY 568 PGAPEDQDDXGGDFSGTGLPDYSAANPIKVTHRCVILENDTVQCDLCLKSLQAWKDH 627  
Db 383 ----- 382  
QY 628 KLHIDHEIETLQNKIKNLREVGRHLKKRPEBCDCHKISYHTQHKRLKHROSSLHPFRK 687  
Db 383 -----DPTCPSL----- 390  
QY 686 GLQEKDKVLLREQKRLKLLKRLQNNDTCSMEGLT- CFTHDNQHWQTAPFWTLGPF 746  
Db 391 -----PGVSQCF-----PD 399  
QY 747 CACTSANNNTYWCMTINETHNLFCEP--ATGFLYFDLNTDPYQLMNAVNTLDRDLN 804  
Db 400 CVCEDAYNTYACVRTMSALWNLQYCEFFDQEVFEVYNLTADPDQITNIAKTIDPELIG 459  
QY 805 QLVOLMELRSCG 818  
Db 460 KMYRLMLQSCG 473

Search completed: February 14, 2004, 22:59:55  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 22:33:50 ; Search time 45 Seconds  
(without alignments)  
1859.261 Million cell updates/sec

Title: US-10-025-966A-6  
Perfect score: 4750  
Sequence: 1 MGPSLVLCILSATVFSLLG.....PEMKRPSKSLGQLWEGWG 870  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1584	33.3	709	2	T16584	hypothetical prote
2	858.5	18.1	552	1	KJHUGU	N-acetylglucosamin
3	452.5	9.5	649	2	S43229	arylsulfatase (EC
4	448.5	9.4	639	2	T46577	arylsulfatase (EC
5	341.5	7.2	646	1	KJXW	arylsulfatase (EC
6	302.5	6.4	497	2	G65169	probable sulfatase
7	301.5	6.3	497	2	C91206	probable sulfatase
8	301.5	6.3	497	2	E86052	probable sulfatase
9	290	6.1	535	2	AF0103	probable sulfatase
10	251	5.3	571	2	G90891	probable sulfatase
11	251	5.3	571	2	E64903	arylsulfatase homo
12	250.5	5.3	557	2	AF0370	probable sulfatase
13	250	5.3	571	2	B85726	probable sulfatase
14	241.5	5.1	517	2	AB0102	probable sulfatase
15	238.5	5.0	583	1	KJHUAC	steryl-sulfatase (
16	238	5.0	551	2	S07089	arylsulfatase (EC
17	238	5.0	567	2	A37362	arylsulfatase (EC
18	234.5	4.9	550	1	KJHUID	iduronate-2-sulfat
19	233	4.9	497	2	AF0506	probable secreted
20	232	4.9	551	2	S01793	arylsulfatase (EC
21	228	4.8	538	2	F83354	probable sulfatase
22	226	4.8	563	2	A47153	iduronate-2-sulfat
23	224	4.7	465	2	F70837	probable sulfatase
24	223.5	4.7	479	2	C75099	hypothetical prote
25	222	4.7	503	2	E83642	choline sulfatase
26	221	4.7	514	2	T44602	phosphonate monoes
27	219	4.6	787	2	B70643	probable sulfatase
28	218.5	4.6	522	1	KJHUG6	N-acetylglactosam
29	211.5	4.5	533	1	KJHUAB	N-acetylglactosam

ALIGNMENTS

RESULT 1

T16584

hypothetical protein K09C4.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16584

R/DU, Z.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid K09C4.

A:Reference number: Z18542

A:Accession: T16584

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-709 <DUZ>

A:Cross-references: EMBL:U43375; NID:g1125841; PID:g1125842; PIDN:AAA83618.1; CESP:K09C

C:Genetics:

A:Gene: CESP:K09C4.8

A:Introns: 33/2; 72/2; 110/2; 180/3; 237/2; 287/3; 346/2; 406/2; 448/3; 545/2; 564/2; 6

Query Match 33.3%; Score 1584; DB 2; Length 709;

Best Local Similarity 40.1%; Pred. No. 1.7e-106;

Matches 316; Conservative 118; Mismatches 197; Indels 158; Gaps 17;

QY	45	NIILVITDDQDVELGSMQVWKNTRIMEQGAHFNAFVTTMCCPSSRSIIITGKYVNH 104
DB	36	NIILVITDDQDIELGSMDFMPTKSQIMKRGFTSGYVTTICCPSSRTILTGLVNH 95
QY	105	NTVTNENSCSSPQWQHESRTFAVINSTGRTAPFGKILNEYNGSYPPQWKEWGLL 164
DB	96	HVHTNNQNTGVWEKRWHEKKSIGVYLOEAGYRTAYLGLKYLNEYDGSYTPPGWDEHAI 155
QY	165	KNSRFNYITLCRNGVKEKHSYSDYLTDLTNDVSFRTSKMYPHRYPVLMVISHAA 224
DB	156	KNSKFNYITWNSGERKEGSEYKDYFDLITNLSKFDIKHAKIRAWQPPALISYPA 215
QY	225	PHGPEDSAPQYGRLPFNPAHQITPSPYNAPNDPKHMYRTGPMKPIHMEFTNMQRKL 284
DB	216	PHGPEDPAPQFAMPENISHTGSWNPAPNDPKQWLLQRTGKMDVHISFTDLLHRL 275
QY	285	QTILMSVDDSMETYNMLVETGLDNTYIVYTADHGVHIGQFGLVKGSKMPYFDIRVPY 344
DB	276	QTILQSVDEGIERLNLRLNLQWNTYAITSDHGHILGQFGLLKGKMPYFDIRVPFF 335
QY	345	VRGPNVEAGLPHPIVYINIDTIDTIDTIDTIDTIDTIDTIDTIDTIDTIDTIDTID 404
DB	336	MRGPGIPRNVTFNEIVTNVDIAPTMLHIAGVKPKARMNGRSLLELVALKKK-----KKK 389
QY	405	----MRVWDSFLVERGKLLH-KRNDKVDAGEENFLPKYQVQKLCQBAEYQACEQIG 459
DB	390	HMTALPWRDITLIERGKMPKLKIRDRYIKQKXF-NKENLSKECKRKRQWQCVH-G 447
QY	460	QKQVCVEDATGKLHLHKKGPMRLGSSRLSNLVPKYQGSGEACTCDSGDYKLSLAGSR 519

Db 448 QLVKCYTVEDRWRYKCR-----DNWSDQSC-----R 476  
QY 520 KKLKKYKASVRSRSVAIEVDGRVYHVLGDAAQPRNLTKRHWPAGPEDDKDG 579  
Db 477 KK-----REI-----SNYDDDDI 489  
QY 580 GDFSGTGLPDYSAANPIKVTTHRCVILENDTVQCDLDLYKSLQAWKDKLHIDHEIETLQ 639  
Db 490 DEF-----LTYADRENFSEGHYQCFEDSGEV---G 519  
QY 640 NKIKNLREVRGHLKKRPBECCHKISYHTQHKGLKRGSLHPFRKGLQEKDKVLLR 699  
Db 520 EBLDGRHSRXLGSK-----CSCSR-----NVSHPIK-----LLE 549  
QY 700 EOKRKKLKLRLKRLQND---TCSMPLGTCFTHDNQHOTAPFW---TLGPFCACTSAN 753  
Db 550 QXMSKXLYKXKQNGSLKPKDLSLPMQNCFTTASHWKTPPLWPELGEFECQNCN 609  
QY 754 NNTYCMRTINETHNPLFCFATGFLFVLDNTDPIQLMNAVNTLDRDLVNLQHLVLMEL 813  
Db 610 NNTYCLRTKNTNETHNPLFCEFTERTISFYDNTDPIQLINAVSLDGLVLEQLSEQLRL 669  
QY 814 RSCGYKQC 822  
Db 670 RKCKN-RQC 677

RESULT 2  
KJHUGU  
N;Alternate names: chondroitinsulfatase; N-acetyl-D-glucosamine-6-sulfate 6-sulfohydrolase  
C;Species: Homo sapiens (man)  
C;Date: 22-Nov-1993 #sequence\_revision 27-Oct-1995 #text\_change 24-Sep-1999  
C;Accession: S27164; A31672  
R;Robertson, D.A.; Freeman, C.; Morris, C.P.; Hopwood, J.J.  
Biochem. J. 288, 539-544, 1992  
A;Title: A cDNA clone for human glucosamine-6-sulphatase reveals differences between aryl  
A;Reference number: S27164; PMID:93098807; PMID:1463457  
A;Accession: S27164  
A;Molecule type: mRNA  
A;Residues: 1-552 <ROB>  
A;Cross-references: GB:212173; EMBL:M23657; NID:G31867; PIDN:CAA78164.1; PID:G31867  
A;Note: parts of this sequence, including the amino end of the mature protein, were determined from a form described with a proteolytic cleavage releasing residue 222 as the amino  
R;Robertson, D.A.; Freeman, C.; Nelson, P.V.; Morris, C.P.; Hopwood, J.J.  
Biochem. Biophys. Res. Commun. 157, 218-224, 1988  
A;Title: Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase.  
A;Reference number: A31672; MUID:89061714; PMID:3196333  
A;Accession: A31672  
A;Molecule type: mRNA  
A;Residues: 178-552 <RO2>  
A;Cross-references: GB:212173; EMBL:M23657; NID:G31866  
A;Gene: GDB:GNS  
A;Cross-references: GDB:120006; OMIM:252940  
A;Map position: 12q14-12q14  
A;Note: defects in this gene can cause mucopolysaccharidosis type III D, Sanfilippo D di  
C;Function:  
A;Description: hydrolyzes N-acetyl-D-glucosamine 6-sulfate units in heparan sulfate and  
C;Superfamily: animal sulfatase  
C;Keywords: glycoprotein; lysosomal storage disease; lysosome; Sanfilippo disease; sulfu  
F;1-43/Domain: signal sequence #status predicted <SIG>  
F;44-552/Product: N-acetylglucosamine-6-sulfatase #status predicted <MAT>  
F;91/Modified site: 3-oxoalanine (Cys) #status predicted  
F;111-117-183-198-210-279-317-362-387-405-449-480/Binding site: carbohydrate (Asn) (cova  
F;422/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 18.1%; Score 858.5; DB 1; Length 552;  
Best Local Similarity 27.2%; Pred. No. 3.9e-54;  
Matches 226; Conservative 76; Mismatches 174; Indels 355; Gaps 18;  
QY 4 PSLVCLLSAT--VFSILGSSAFSLSHRLKGRFQDRNRIPNIIIVLTDDQDVELGSM 61

Db 24 PALLLVLCGLGVGAAGR-----RNVVLLLTDDQDVELGSM 64  
QY 62 QVNNKTRRIMEQGAHFINAFTVPMCCPSRSILTKYVHNHTYTN--NENCSPSQW 119  
Db 65 TPLKTKTALIGEMGMTFSAYYPSALCCPSRASILTKYPPHNHVVNNTLEGCSKSKW 124  
QY 120 AQHESRTFAVILNS-TGYRTAFGKYLANEYNG-----SYVPPGKWKWYGLLKRNRFNY 172  
Db 125 KIOEPNTFAILRSMCGYQTFAGKYLANEYAGDAGLEHVPLGWSYVALEKNSKYNY 184  
QY 173 TLCRNGVKHSDYSKYLTDLITNDSVSPFRTSKMYPHRPVLWVISHAAPHGPESA 232  
Db 185 TLSINGKARKHGENYVDYLTDLVANVSLDFLDYSNF---EPFFMMIATPAPHSPWTA 241  
QY 233 PQYSRLFPNASQHITPSYNAFNPDKHWMRYT-GPMKPIHMEFTNMLQKRLQTLMSYD 291  
Db 242 PQYKAFQVAPPRKNFN-IHGTNKGWILIRAKFTMTNSSIQFLDNAFRKRWQLLSYD 300  
QY 292 DSMETIYNMLVETGELDNTYIVYADHGHVHIQFGLVKGKSMYPFDIRVPPYVPGNVE 351  
Db 301 DLVKLVKRLFTGELNNTYIIFTSDNGYHTQGFSLPIDKRLQYEFDIKVPILLVRGP 360  
QY 352 AGCLNPHIVLINDLAPTILDIAGLDI-PADMDGKSLKLLDTERPVNRPHLKKKQWRD 410  
Db 361 PNQTSKMLVANIDLGTILDIAGYDLNKTQMDMSLLPIL---RGASNL-----TWRS 410  
QY 411 SPLVGRKLLHRRDNDKVDQAEENFLPKYQVVKDLQCRABYQTAACEQLQKQCVEDATG 470  
Db 411 DVLVE----- 415  
QY 471 KKLKHKCGMRGLGSRALSNLVPKYVGSGSEACTCDSDGYKLSLAGRRKLFKKYKAS 530  
Db 416 -----YQEG----- 420  
QY 531 YVRSRSIRSVIAIEVDGRVYHVLGDAAQPRNLTKRHWPAGPEDDKDGDFSGTGLPD 590  
Db 421 -----RNVT----- 424  
QY 591 YSAANPIKVTTHRCVILENDTVQCDLDLYKSLQAWKDKLHIDHEIETLQNKIKNLREVRG 650  
Db 425 ----- 424  
QY 651 HLKKRPEECCHKISYHTQHKGLKRGSLHPFRKGLQEKDKVLLREKREKKLRKL 710  
Db 425 -----DPTCSLS----- 432  
QY 711 LKRLQNNTCSMPGLT-CFTHDNQHWQTAFTWLGPFCACTSANNTYCMRTINETHN 769  
Db 433 -----PGVSQCF-----PDCVCEADAYNNTYACVTRMSALWNL 464  
QY 770 LFCFET-ATGFLEYFDLNTDPPYOLMNAVNTLDRDLVNLQHLVLMELRSCG 818  
Db 465 QYCEFDQEVFVEVYNLTADPDQITIAKTIIDPELLGKNYRLMLQSCSG 515

RESULT 3  
S43229  
A;Title: arylsulfatase (EC 3.1.6.1) precursor type I [validated] - Volvox carteri  
C;Species: Volvox carteri  
C;Date: 28-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 18-Aug-2000  
C;Accession: S43229; S68892; S41433  
R;Hallmann, A.; Sumper, M.  
Eur. J. Biochem. 221, 143-150, 1994  
A;Title: An inducible arylsulfatase of Volvox carteri with properties suitable for a re  
A;Reference number: S43229; MUID:94222055; PMID:8168504  
A;Accession: S43229  
A;Molecule type: mRNA  
A;Residues: 1-649 <HAL>  
A;Cross-references: EMBL:X77214; NID:G452338; PIDN:CAA54426.1; PID:G452339  
A;Note: part of this sequence, including the amino end of the mature protein was confir  
R;Selmer, T.; Hallmann, A.; Schmidt, B.; Sumper, M.; von Figura, K.  
Eur. J. Biochem. 238, 341-345, 1996

A;Title: The evolutionary conservation of a novel protein modification, the conversion of  
A;Reference number: S68892; MUID:96283826; PMID:8681943  
A;Accession: S68892  
A;Molecule type: protein  
A;Residues: 64-71, 'S', 73-76 <SEL>  
C;Description:  
A;Note: Required for mineralization of sulfate  
A;Note: remarkably insensitive towards detergents like dodecyl sulfate  
C;Superfamily: plant sulfatase  
C;Keywords: Glycoprotein; periplasmic space; sulfuric ester hydrolase  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-649/Product: arylsulfatase #status experimental <NAT>  
F;72/Modified site: 3-oxoalanine (Cys) #status experimental

Query Match 9.5%; Score 452.5; DB 2; Length 649;  
Best Local Similarity 27.4%; Pred. No. 1.2e-24;  
Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;  
QY 6 LVCLLSATVFLGSSAFSLSHRLKGRFQDRDRNRIRNIIILVLTDDQDVELGSM--QV 63  
Db 7 VALCLLG---FAALTA---HQ-----RPNFVIFTDDQDQIGNSTHPY 47  
QY 64 MNKTRIMEQGAHFINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSSP-----SW 118  
Db 48 OPKLHEHRYPGIELKNYFVTPVCCPSRTNLWRGQFSHN---TNPTDVLGPHGYAKW 103  
QY 119 QAOHESRTF-AVYLNSTGYRTAFPGKYLNEY---NGSVYPPGKWEVGLLKNSRF--NYIT 173  
Db 104 KSLGIDKSYLPWLNQNGYNTYYGKFLVDYSVSNYQNPVAGTDDIDALVTPYTFDYNP 163  
QY 174 -LCRRNGVKEKGGSDYKDYLTDLITNDSVSPFRSKMYPHRVLVMTISHAAPHGDS 232  
Db 164 GFSRNGATP---NIYPGYSTVDIADKVAQIKTA--VAACKPFYAQISPIAPH----- 212  
QY 233 POYSRLFPNASQHTTPSYNAPNP-DKHWIMRYGPM-----KPIHMEF 275  
Db 213 -TSTQIYDPVANATKTFYPPIPAPRIHELFSDATLPEGTSHKNLYEADVSDKPWIRA 271  
QY 276 TMLQ-----RKRLOTMSVDDSMETIYNMLVETGELONTYVVTADHGHHIGQ 324  
Db 272 LPLAQNRTVLEEVRLRLSLASVDELDIRVATLQEAQVNDITLITSADNGYHVT 331  
QY 325 FGLVKGKSMPEFDIRVPFYVRGPNVEAG----CLNPHIVLINIDLAFTILDIAG----- 374  
Db 332 HRFAGKATYADEDLRVPLIRGPGIRASHSDKPKANSKVLHVDPAFTILTLACAGDQVG 391  
QY 375 -----LDIPADMDGKSLKLLDTERPVNRFLKKQVRWDSFLVE-----RGKLLH 421  
Db 392 DKALDGTGLGLYANDGN---LLADYPRPNH-----RNQFQGEFWGWSDEVHL 438  
QY 422 KRNDKVDQAQENFLPKYQ-----RVKDLQRAEYQTAQEQGKQWCQVEDATGKLL 474  
Db 439 -----HIPRYTNNSKAVRYD-----EDNQAKKLIVSTCNEREL 474  
QY 475 HKCK 478  
Db 475 YDLK 478

## RESULT 4

T46577  
arylsulfatase (EC 3.1.6.1) [validated] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 04-Mar-2000  
C;Accession: T46577  
R;Baker, D.L.; Paletta, J.V.  
submitted to the EMBL Data Library, February 1997  
A;Description: Molecular characterization of the arylsulfatase gene of Neurospora crassa  
A;Reference number: Z23090  
A;Accession: T46577  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-639 <BAK>  
A;Cross-references: EMBL:U89492; PIDN:AA02716.1  
A;Experimental source: strain wild type 74-OR23-1A  
C;Genetics:  
A;Gene: ars-1  
A;Map position: 7  
A;Introns: 115/1, 518/3  
C;Function:  
A;Description: EC 3.1.6.1 [validated, MUID:89384589]  
A;Note: genes are expressed under conditions of sulfur limitation and are under coordin  
C;Superfamily: plant sulfatase  
C;Keywords: sulfuric ester hydrolase

Query Match 9.4%; Score 448.5; DB 2; Length 639;  
Best Local Similarity 27.5%; Pred. No. 2.3e-24;  
Matches 137; Conservative 73; Mismatches 212; Indels 77; Gaps 14;  
QY 13 ATVFSLLGSSAFSLSHRLKGRFQDRDRNRIRNIIILVLTDDQDVELGSMVQNKRRIME 72  
Db 15 STAFVDSQKSLGFEHKPKGTTTEGKK-SNIVFILTDDQDLHLQSLDYLPLKKYLA 73  
QY 73 QGAHFINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSSPSWQAQHSRTF-AVY 130  
Db 74 DEGTYYKRYCTTAICCPARVSLWTGQAHTNNTVDVSPYGGYKPFISQGFNEAYLPW 133  
QY 131 LKSTGYRTAFPGKYLNEY---EYNGSVYPPGKWEVGLLK--NSRFYNYTLCRNGVKEKH 183  
Db 134 LQKAGDITTYTQKLFNAHTVDNYDSFYI-AGWNGSDFLLDPYTSYLNATFORN---RDP 189  
QY 184 GSDYSKDYLTDLITNDSVSPFRSKMYPHRVLVMTISHAAPHG----- 227  
Db 190 PISYEQYSVDVLAEKAYGLDELAANKVHNRRPFLGIAPHSNVEPGFPPSSSSSSSS 249  
QY 228 -----PEDSAPQYGRFLFPNASQHTTPSYNAPNP-DKHWIMR 263  
Db 250 DSATLHRRRTNEHDDIEKSVSFTPIPAARHAHLFPDIVPRTPHFNFSRSGASVSWIR 309  
QY 264 YTGPMKPIEMETNMLQKRLQTLMSVDDSMETIYNMLVETGELONTYVVTADHGHHIG 323  
Db 310 LP-HQSAENVAENDFYRQLRALESVDLVAGVVARLERHGLENTYVVTADNGYHIG 368  
QY 324 QGLVKGKSMPEFDIRVPFYVRGPNVEAGCLNPHIVLINIDLAFTILDIAG----- 375  
Db 369 QHRLQPGKECGFEEDINVLPIVRGPGVKKEVAEVVTSHTVDLAPTVLKLAGALGDKESKE 428  
QY 376 ---DIPADMDGKSI---LKLLDTERPVNRFLKKQVRWDSFLVERGKLLHKNRDNKVD 429  
Db 429 EEEEREYGLDGEAIPVTAEADLAQAKAKGRHEHVTVEYW--GFAVSEGRVPEK-----D 480  
QY 430 AQEENFLPKYQVRKDLQCOR 448  
Db 481 ATRYTNNTYKAVKILGER 499

## RESULT 5

KJKN  
arylsulfatase (EC 3.1.6.1) precursor - Chlamydomonas reinhardtii  
N;Alternate names: sulfatase  
C;Species: Chlamydomonas reinhardtii  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 24-Sep-1999  
C;Accession: JQ0310  
R;de Hostos, E.L.; Schilling, J.; Grossman, A.R.  
Mol. Gen. Genet. 218, 229-239, 1989  
A;Title: Structure and expression of the gene encoding the periplasmic arylsulfatase of  
A;Reference number: JQ0310; MUID:89384447; PMID:2476654  
A;Accession: JQ0310  
A;Molecule type: mRNA  
A;Residues: 1-646 <DEH>  
A;Cross-references: GB:X16180; NID:g18118; PIDN:CAA34302.1; PID:g18119; GB:X52304; NID:g  
A;Experimental source: strain cw15mt+  
A;Note: part of this sequence, including the amino end of the mature protein, was determ  
C;Comment: This enzyme is commonly produced by soil microorganisms and plays an importar  
C;Superfamily: plant sulfatase







A:Gene: ECs2103  
C:Superfamily: animal sulfatase

```
Query Match          5.3%; Score 251; DB 2; Length 571;
Best Local Similarity 22.7%; Pred. No. 3.9e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

QY 18 LLGSSAFSLSHRLKGRFQDRRNI-----RPNIILVLTDDQ-----54
Db 29 LASGMAFAAADDVVKLKATKTNVAFSDFTPTTEYSTKGKRNIIIVLTMDDLGYQLPFDK 88
QY 55 -----DVLGSMQVNNKTRRIMEQGAHFNAFVTTMCCPSRS 93
Db 89 GSFPDKTNENREVDVTKIGIDKAIEAAQKSTPTLLSLMDEGVRFRTGYVAHGVSGPSRA 148
QY 94 SILTKGVVHNNTYNNENSCSSPSWQAQHE---SRTEAVYL-NSTGYRTAFPGK-VLYNE 148
Db 149 AIMGTRAPARGVYSNTD-----AODGIPLTETPLPQLFQNHGYTYAAGKWLHLSKI 200
QY 149 NGSVYPPGKKEWGLLKNRPFY--NYTL--CRNGVKKKGSDY-----187
Db 201 SNVPVPED-----KQTRDYHDNFTTSAEEWQPNRGDFYFMGFHAGTAYNPSPL 252
QY 188 -----SKDYLTLTNDVSFRTSKMYPRPVLVISHAAPHGPDSS-AP-OYSR 237
Db 253 FKNERVPAKGYISDQLTDEAIGVDRAKTL--DQPMVLVLYANAPHLNDNPAPDQYOK 310
QY 238 LFPNASQHIPTSYNAPNDPKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETI 297
Db 311 QFNTGSQ--TADNYTA-----SVYSVDQGVKRI 336
QY 298 YNMLVETGELDNTVIYVTAADGVHI--GQF---GLVKG-KSMPEYEDIRVPPYVGRPNVEA 352
Db 337 LEQKKKGQVNTIILFTSNGAVIDGLPLNGAQKGYKSTPGGTHTPMFMW-----K 392
QY 353 GCLNP-----HIVLNIDLAFTILDIAGLDIPAD--MDGKSILKLLDTER---PVNRFHLKK 403
Db 393 GKLPQGNVYKLIISANDFPVPTALDAADISIPKDLKLGVSLLPWLQDKKQGEPHKNLTWIT 452
QY 404 KQVWRDSFLVERGKLAHK-----RDND--KYDAQENFL 436
Db 453 SYSWFDEENIPFDWNTVHKFVRHOSDDYPHNPTEDLSQFSYTVRNNDYSLVYTVENNQL 512
QY 437 PKYQVRKDLQARAEYQACQLGQKQCV 465
Db 513 GLY-KLTDLQCKNLAANPQVVKMGV 540

RESULT 11
B64903
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64903
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-571 <BLAT>
A:Cross-references: GB:AE000247; GB:U00096; NID:g1787773; PIDN:AAC74571.1; PID:g1787775;
A:Experimental source: strain K-12, substrain MG1685
C:Superfamily: animal sulfatase

Query Match          5.3%; Score 251; DB 2; Length 571;
Best Local Similarity 22.7%; Pred. No. 3.9e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

QY 18 LLGSSAFSLSHRLKGRFQDRRNI-----RPNIILVLTDDQ-----54
Db 29 LASGMAFAAADDVVKLKATKTNVAFSDFTPTTEYSTKGKRNIIIVLTMDDLGYQLPFDK 88
QY 55 -----DVLGSMQVNNKTRRIMEQGAHFNAFVTTMCCPSRS 93
Db 89 GSFPDKTNENREVDVTKIGIDKAIEAAQKSTPTLLSLMDEGVRFRTGYVAHGVSGPSRA 148
QY 94 SILTKGVVHNNTYNNENSCSSPSWQAQHE---SRTEAVYL-NSTGYRTAFPGK-VLYNE 148
Db 149 AIMGTRAPARGVYSNTD-----AODGIPLTETPLPQLFQNHGYTYAAGKWLHLSKI 200
QY 149 NGSVYPPGKKEWGLLKNRPFY--NYTL--CRNGVKKKGSDY-----187
Db 201 SNVPVPED-----KQTRDYHDNFTTSAEEWQPNRGDFYFMGFHAGTAYNPSPL 252
QY 188 -----SKDYLTLTNDVSFRTSKMYPRPVLVISHAAPHGPDSS-AP-OYSR 237
Db 253 FKNERVPAKGYISDQLTDEAIGVDRAKTL--DQPMVLVLYANAPHLNDNPAPDQYOK 310
QY 238 LFPNASQHIPTSYNAPNDPKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETI 297
Db 311 QFNTGSQ--TADNYTA-----SVYSVDQGVKRI 336
QY 298 YNMLVETGELDNTVIYVTAADGVHI--GQF---GLVKG-KSMPEYEDIRVPPYVGRPNVEA 352
Db 337 LEQKKKGQVNTIILFTSNGAVIDGLPLNGAQKGYKSTPGGTHTPMFMW-----K 392
QY 353 GCLNP-----HIVLNIDLAFTILDIAGLDIPAD--MDGKSILKLLDTER---PVNRFHLKK 403
Db 393 GKLPQGNVYKLIISANDFPVPTALDAADISIPKDLKLGVSLLPWLQDKKQGEPHKNLTWIT 452
QY 404 KQVWRDSFLVERGKLAHK-----RDND--KYDAQENFL 436
Db 453 SYSWFDEENIPFDWNTVHKFVRHOSDDYPHNPTEDLSQFSYTVRNNDYSLVYTVENNQL 512
QY 437 PKYQVRKDLQARAEYQACQLGQKQCV 465
Db 513 GLY-KLTDLQCKNLAANPQVVKMGV 540

RESULT 12
AF0370
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-557 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92289.1; PID:g15981000; GSPDB:GN00175
C:Genetics:
C:Superfamily: animal sulfatase
C:Keywords: sulfuric ester hydrolase

Query Match          5.3%; Score 250.5; DB 2; Length 557;
Best Local Similarity 21.3%; Pred. No. 4.1e-10;
Matches 114; Conservative 88; Mismatches 171; Indels 161; Gaps 27;

QY 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNIILVLT 51
Db 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNIILVLT 51
QY 52 D-----DQDV-----ELGSMQVNNKTR-----IMEQGAHFNA 80
Db 66 DDLGVQLPDKTSFDPKSMEDRDVDTYKIGIDKAIEAAKSTPTLLSLMDE-GVRLTN 124
```

```
Db 29 LASGMAFAAADDVVKLKATKTNVAFSDFTPTTEYSTKGKRNIIIVLTMDDLGYQLPFDK 88
QY 55 -----DVLGSMQVNNKTRRIMEQGAHFNAFVTTMCCPSRS 93
Db 89 GSFPDKTNENREVDVTKIGIDKAIEAAQKSTPTLLSLMDEGVRFRTGYVAHGVSGPSRA 148
QY 94 SILTKGVVHNNTYNNENSCSSPSWQAQHE---SRTEAVYL-NSTGYRTAFPGK-VLYNE 148
Db 149 AIMGTRAPARGVYSNTD-----AODGIPLTETPLPQLFQNHGYTYAAGKWLHLSKI 200
QY 149 NGSVYPPGKKEWGLLKNRPFY--NYTL--CRNGVKKKGSDY-----187
Db 201 SNVPVPED-----KQTRDYHDNFTTSAEEWQPNRGDFYFMGFHAGTAYNPSPL 252
QY 188 -----SKDYLTLTNDVSFRTSKMYPRPVLVISHAAPHGPDSS-AP-OYSR 237
Db 253 FKNERVPAKGYISDQLTDEAIGVDRAKTL--DQPMVLVLYANAPHLNDNPAPDQYOK 310
QY 238 LFPNASQHIPTSYNAPNDPKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETI 297
Db 311 QFNTGSQ--TADNYTA-----SVYSVDQGVKRI 336
QY 298 YNMLVETGELDNTVIYVTAADGVHI--GQF---GLVKG-KSMPEYEDIRVPPYVGRPNVEA 352
Db 337 LEQKKKGQVNTIILFTSNGAVIDGLPLNGAQKGYKSTPGGTHTPMFMW-----K 392
QY 353 GCLNP-----HIVLNIDLAFTILDIAGLDIPAD--MDGKSILKLLDTER---PVNRFHLKK 403
Db 393 GKLPQGNVYKLIISANDFPVPTALDAADISIPKDLKLGVSLLPWLQDKKQGEPHKNLTWIT 452
QY 404 KQVWRDSFLVERGKLAHK-----RDND--KYDAQENFL 436
Db 453 SYSWFDEENIPFDWNTVHKFVRHOSDDYPHNPTEDLSQFSYTVRNNDYSLVYTVENNQL 512
QY 437 PKYQVRKDLQARAEYQACQLGQKQCV 465
Db 513 GLY-KLTDLQCKNLAANPQVVKMGV 540

RESULT 12
AF0370
A:Title: Genome sequence of Yersinia pestis (strain CO92)
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-557 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92289.1; PID:g15981000; GSPDB:GN00175
C:Genetics:
C:Superfamily: animal sulfatase
C:Keywords: sulfuric ester hydrolase

Query Match          5.3%; Score 250.5; DB 2; Length 557;
Best Local Similarity 21.3%; Pred. No. 4.1e-10;
Matches 114; Conservative 88; Mismatches 171; Indels 161; Gaps 27;

QY 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNIILVLT 51
Db 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNIILVLT 51
QY 52 D-----DQDV-----ELGSMQVNNKTR-----IMEQGAHFNA 80
Db 66 DDLGVQLPDKTSFDPKSMEDRDVDTYKIGIDKAIEAAKSTPTLLSLMDE-GVRLTN 124
```



QY 391 --DTERFVNRHL-----KKQKRVWRDSEFLVER--GK-----LLHKRDNDKVDQAE 432  
Db 416 RPKGDKRPTSLQYTMPYGGQSYGQGVTRDYTLVIDRKVGKPLTYLTHDNKNDPQMKN 475  
QY 433 ---ENFLPKYQVXDLQORABYQYACQOLGQKQVQVEDATGKLKJH 475  
Db 476 IAAENMALVNQIAD-----ELIPWLEHSGDVWRPTVEVAANAAY 516

RESULT 15  
KTHUAC  
steryl-sulfatase (EC 3.1.6.2) precursor - human  
N;Alternate names: arylsulfatase C; steroid sulfatase (STS); steryl-sulfate sulfohydrolase  
C;Species: Homo sapiens (man)  
C;Date: 21-May-1990 #sequence revision 27-Oct-1995 #text change 11-May-2000  
C;Accession: A32641; A57116; A25961; S05415; S05423; I52800; I65619  
J;Stein, C.; Hille, A.; Seidel, J.; Rijnhout, S.; Waheed, A.; Schmidt, B.; Geuze, H.; van  
J. Biol. Chem. 264, 13865-13872, 1989  
A;Title: Cloning and expression of human steroid-sulfatase. Membrane topology, glycosylation  
A;Reference number: A32641; MUID:89340479; PMID:2668275  
A;Accession: A32641  
A;Molecule type: mRNA  
A;Residues: 1-583 <STE>  
A;Cross-references: GB:J04964; NID:G338564; PIDN:AAA60597.1; PID:G338565  
A;Experimental source: BHK-21 cells  
A;Note: Parts of this sequence were determined by protein sequencing  
R;Yen, P.H.  
unpublished results 1988, cited by GenBank  
A;Reference number: A57116  
A;Accession: A57116  
A;Molecule type: mRNA  
A;Residues: 1-22, 'E', 24-583 <YEN1>  
A;Cross-references: GB:M16505; NID:G338513; PIDN:AAA60596.1; PID:G338514  
R;Yen, P.H.; Allen, E.; Marsh, B.; Mohandas, T.; Wang, N.; Taggart, R.T.; Shapiro, L.J.  
Cell 49, 443-454, 1987  
A;Title: Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of  
A;Reference number: A25961; MUID:87187642; PMID:3032454  
A;Accession: A25961  
A;Molecule type: mRNA  
A;Residues: 1-22, 'E', 24-456, 'LRTTHPSGRFPSSPTSTFWFORIALPHTCASVGVMSPTTHLYSLFF', 'KIPERE  
A;Cross-references: GB:M16505; NID:G338513  
A;Note: This sequence revised in A57116  
R;Kawano, J.I.; Kotani, T.; Ohkaki, S.; Minamino, N.; Matsuo, H.; Oinuma, T.; Aikawa, E.  
Biochim. Biophys. Acta 997, 199-205, 1989  
A;Title: Characterization of rat and human steroid sulfatases.  
A;Reference number: S05414; MUID:89352671; PMID:2765556  
A;Accession: S05415  
A;Molecule type: protein  
A;Residues: 22-43, 'X', 45-46 <KAW>  
R;Dibbelt, L.; Otto, J.; Kues, E.  
Biol. Chem. Hoppe-Seyler 370, 847-848, 1989  
A;Title: The N-terminal amino-acid sequence of human placental steryl-sulfatase.  
A;Reference number: S05423; MUID:90074181; PMID:2590467  
A;Accession: S05423  
A;Molecule type: protein  
A;Residues: 22-35 <DiB>  
R;Yen, P.H.; Marsh, B.; Allen, E.; Teai, S.P.; Ellison, J.; Connolly, L.; Neiswanger, K.  
Cell 55, 1123-1135, 1989  
A;Title: The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence  
A;Reference number: I52800; MUID:89077541; PMID:3203382  
A;Accession: I52800  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 134-274 <YEN3>  
A;Cross-references: GB:M23945; NID:G338604; PIDN:AAA60598.1; PID:G338607  
A;Accession: I65619  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 461-583 <YEN4>  
A;Cross-references: GB:M23556; NID:G338605; PIDN:AAA60599.1; PID:G338608  
C;Genetics:

A;Gene: GDB:STS  
A;Cross-references: GDB:120393; OMIM:308100  
A;Map position: Xp22.32-Xp22.32  
A;Note: defects in this gene can cause X-linked ichthyosis  
C;Function:  
A;Description: hydrolyzes 3beta-hydroxysteroid sulfates to release sulfate  
C;Superfamily: animal sulfatase  
C;Keywords: endoplasmic reticulum; glycoprotein; lysosome; microsomal; sulfuric ester h  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-583/Product: steryl-sulfatase #status predicted <MAT>  
F;185-211/Domain: transmembrane #status predicted <TM1>  
F;213-237/Domain: transmembrane #status predicted <TM2>  
F;47,259/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;75/Modified site: 3-oxoalanine (Cys) #status predicted  
F;333,459/Binding site: carbohydrate (Asn) (covalent) #status absent  
Query Match 5.0%; Score 238.5; DB 1; Length 583;  
Best Local Similarity 20.7%; Pred. No. 3.2e-09; Indels 237; Gaps 23;  
Matches 126; Conservative 67; Mismatches 180;  
QY 43 RPNILVLTDDQVELGSMQVM-NKTRRI-----MEQGAHFNAFVTPMCCPSRSIL 96  
Db 26 RPNILVMAD--DLGIGDPCGCGKNTIRTNIDRLASGGVKLTQHLAASPLCTPSRAAFM 83  
QY 97 TGYVHNHTYNNENCSSPSWQ-----AQHESRTPAVYLNSTGYTAFFGK 143  
Db 84 TGRY-----PVRSGMASWSTGVFLFTASSGGLPTDEITFAKLLKQGYSTALIGK 134  
QY 144 Y-----LNEY-----NGSYVPPGWKE----- 159  
Db 135 WHLGWSCHSKTDFCHPLHGHFNFYGISTNLNRDCKPGEVSFTTGFKLVFLQIVG 194  
QY 160 -----WVGLLNSRFNYTLCRNGVKEHG 184  
Db 195 VTLTLAALNCLGLHLVPLGVFFSLLFLAALILTLFLGFLHYRPLNCFMRN--YBIIQ 252  
QY 185 SDYSKDYLTDLITNDVSFFRTSKMYPHPVLMVISHAAPHGPEDSAPQYSLFFNVAQ 244  
Db 253 QPMYSDNLTQRLTVEAAQFQRTVE-----TPFLVLVSLVHV-----TALPSSKDFACKSQ 304  
QY 245 HTSPYNAYPNPKHWIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNMLVET 304  
Db 305 K--GVYGD-----VEEMDWSVQIILNLDL 329  
QY 305 GELDNTYIVYTDHGYHI-----GQFGLVK-GKSNMPEFDIRVPPYVGGKPY-E 351  
Db 330 RLANDTLIYFTSDQGAHVEVSVSKGEIHGGNGIYKGGKANNWEGGIRVPGILRWPRVQ 389  
QY 352 AGCLNPHIVLNIADLPTIDTAGLDIPAD--MDGKSILKLLD--TERPVN----- 397  
Db 390 AGQKIDEPTSNMIDFFPTVAKLAGAPLPEDRIIDGRDLMLLEGKSQSDHEFLPHYCNAY 449  
QY 398 ----RPHLKKQRYWRDSFLVER-----GKLLHKRD-----NDKVIDA 430  
Db 450 LNAVVRHPQNSTSIWKAFFFTFNFNPNVGSNGCFATHVCFGFSYVTHHDPFLFLDISKDP 509  
QY 431 QEENFL-----PKYQVXDLQORAE-----YQTACEQLGQK 461  
Db 510 RERNPLTPASEPRFYELKWKQEAADRHQTLPPEVPPDQFSWNNFLWKPWLQLCCPSTGLS 569  
QY 462 WQVEDATGK 471  
Db 570 CQCDREKQDK 579

Search completed: February 14, 2004, 22:58:59  
Job time : 50 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 19:39:30 ; Search time 70 Seconds  
(without alignments)  
1972.745 Million cell updates/sec

Title: US-10-025-966A-6

Perfect score: 4750

Sequence: 1 MGPPSLVCLLSATVFSLLG.....PEMKRPSKSLGQLWEGWG 870

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Genesec 19Jun03.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4750	100.0	870	22	Human drug metabol
2	4750	100.0	870	22	Human protein SEQ
3	4750	100.0	870	22	Human gene 5' end
4	4750	100.0	870	23	Human albumin fusi
5	4750	100.0	870	23	Human albumin fusi
6	4746	99.9	870	22	Human gene 5' end
7	4746	99.9	870	23	Human sulphatase p
8	4746	99.9	870	23	Human albumin fusi
9	4746	99.9	870	23	Human polypeptide

10	4746	99.9	885	23	ABB80922	Human KIAA1427 pol
11	4723.5	99.4	867	21	AAV66648	Membrane-bound pro
12	4723.5	99.4	867	22	AAU12181	Human PRO1120 poly
13	4723.5	99.4	867	22	AAU12181	Human PRO1120 poly
14	4723.5	99.4	867	23	ABB95459	Human angiotensin
15	4723.5	99.4	867	23	ABB94853	Human PRO1120 prot
16	4723.5	99.4	867	24	ABU66579	Human PRO polypept
17	4723.5	99.4	867	24	ABU66585	Human secreted/tr
18	4723.5	99.4	867	24	ABU59660	Novel secreted and
19	4723.5	99.4	867	24	ABU53064	Novel human secret
20	4723.5	99.4	867	24	ABU53211	Human secreted/tr
21	4723.5	99.4	867	24	ABU59360	Novel human secret
22	4723.5	99.4	867	24	ABU60495	Human secreted/tr
23	4723.5	99.4	867	24	ABU57986	Human PRO polypept
24	4723.5	99.4	867	24	ABU58917	Human secreted/tr
25	4723.5	99.4	867	24	ABU13877	Human PRO1120 poly
26	4723.5	99.4	867	24	ABU10832	Human PRO polypept
27	4616	97.2	850	22	AAW79216	Human protein SEQ
28	4473.5	94.2	875	23	ABG71838	Mouse SULF2 protei
29	4338	91.3	818	21	AAW42268	Human ORFX ORF2032
30	3524	74.2	643	22	AAW80199	Human protein SEQ
31	3524	74.2	643	22	AAW80200	Human protein SEQ
32	3179.5	66.9	871	21	AAW00191	Breast cancer prot
33	3179.5	66.9	871	22	AAW85481	Human SULF1 sulfat
34	3179.5	66.9	871	23	ABG71835	Human secreted pro
35	3179.5	66.9	871	23	ABG69640	Breast cancer asso
36	3179.5	66.9	871	24	ABR47506	Quail sulfatase (Q
37	3121.5	65.7	867	22	AAE00434	Lung small cell ca
38	3063	64.5	818	23	AAU69417	Human sulfatase (H
39	3063	64.5	1611	22	AAE00438	Human protein sequ
40	2977	62.7	800	22	AAW25714	Human polypeptide,
41	2696	56.8	490	22	AAW93919	Human polypeptide,
42	2592	56.7	490	22	AAW93864	Human gene 5' end
43	2356	49.6	434	22	AAE01522	Mouse sulfatase (M
44	2093	44.1	455	22	AAE00437	Drosophila melanog
45	1718.5	36.2	1114	22	ABB71505	

#### ALIGNMENTS

#### RESULT 1

AAW85774  
ID AAW85774 standard; Protein; 870 AA.  
AC AAW85774;  
XX  
XX 29-OCT-2001 (first entry)  
DT  
XX Human drug metabolizing enzyme (ID No. 1558210CD1).  
XX  
XX Drug metabolizing enzyme; DME; immunosuppressive; cytostatic; ophthalmic;  
XX hepatotropic; antiallergic; antiasthmatic; antibacterial; antiviral;  
XX antisense therapy; gene therapy; human.  
XX  
XX Homo sapiens.  
XX  
XX WO200159127-A2.  
XX  
XX 16-AUG-2001.  
XX  
XX 08-FEB-2001; 2001WO-US04423.  
XX  
XX 11-FEB-2000; 2000US-0181856.  
XX 17-FEB-2000; 2000US-0183684.  
XX 25-FEB-2000; 2000US-0185141.  
XX 03-MAR-2000; 2000US-0186818.  
XX 09-MAR-2000; 2000US-0188345.  
XX 17-MAR-2000; 2000US-0189997.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lal P;  
PI



PI Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JL;  
XX WPI; 2001-514673/56.  
DR N-PSDB; AAH76201.  
XX  
XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the  
PT diagnosis, treatment, and prevention of autoimmune/inflammatory, cell  
PT proliferative, developmental and endocrine disorders -  
XX  
XX Claim 1; Page 133-134; 150pp; English.  
XX  
XX The invention provides human drug metabolizing enzymes (DME) and  
CC polynucleotides encoding the DMEs. The DME can be expressed by standard  
CC recombinant methodology. DMEs and their agonists and antagonists are  
CC useful for the diagnosis, treatment, and prevention of autoimmune/  
CC inflammatory, cell proliferative, developmental, endocrine such as  
CC aneurysm, eye, metabolic, and gastrointestinal disorders, including liver  
CC disorders and infection. The present sequence represents a human DME.  
XX  
SQ Sequence 870 AA;  
  
Query Match 100.0%; Score 4750; DB 22; Length 870;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGPPSLVLCILSATVPSLLGGSSAPLSHRLKGRFORDRNIRNIIILVLTDDQDVELGS 60  
Db 1 MGPPSLVLCILSATVPSLLGGSSAPLSHRLKGRFORDRNIRNIIILVLTDDQDVELGS 60  
  
QY 61 MQVMNKRIRMEQGAHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSSQWA 120  
Db 61 MQVMNKRIRMEQGAHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSSQWA 120  
  
QY 121 QHSETFAYVNSTGYRTAFPGKYLNEVNGSVPPGKWEVGLLNKSRFYNYTLCRNGVK 180  
Db 121 QHSETFAYVNSTGYRTAFPGKYLNEVNGSVPPGKWEVGLLNKSRFYNYTLCRNGVK 180  
  
QY 181 EKHGSDYKDYLTDLITNDSSVFFRTSKMTYPRPVLVMSHAAPHGPEDSAPOYSRLFP 240  
Db 181 EKHGSDYKDYLTDLITNDSSVFFRTSKMTYPRPVLVMSHAAPHGPEDSAPOYSRLFP 240  
  
QY 241 NASQHTPSYNAPNPDKHWIMRYTGMKPIHMEFTNMQRLQTLMSVDDSMETIYNN 300  
Db 241 NASQHTPSYNAPNPDKHWIMRYTGMKPIHMEFTNMQRLQTLMSVDDSMETIYNN 300  
  
QY 301 LYETGELDNITYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPYVGNPVEAGCLNPHIV 360  
Db 301 LYETGELDNITYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPYVGNPVEAGCLNPHIV 360  
  
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNFKLKKQWVRNDSFLVERGKLL 420  
Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNFKLKKQWVRNDSFLVERGKLL 420  
  
QY 421 HKRDNDKVDQAEENFLPKYQVVKDLCQAEYQTAQCEQLGQWQVDEATGKLLHKCKGP 480  
Db 421 HKRDNDKVDQAEENFLPKYQVVKDLCQAEYQTAQCEQLGQWQVDEATGKLLHKCKGP 480  
  
QY 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLKSLAGRRKLLFKKTKYKASVYRSRSIRSV 540  
Db 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLKSLAGRRKLLFKKTKYKASVYRSRSIRSV 540  
  
QY 541 ALEVDGRVYHVLGDAAPRNLTKEHWPAGPDDQDDGDFSGTGGLPDYSAANPIKVT 600  
Db 541 ALEVDGRVYHVLGDAAPRNLTKEHWPAGPDDQDDGDFSGTGGLPDYSAANPIKVT 600  
  
QY 601 HRCYILENDTVQCDLDLKYSLQAWDKHLHDHETLQNKIKNLRVGRHGLKKRPEEC 660  
Db 601 HRCYILENDTVQCDLDLKYSLQAWDKHLHDHETLQNKIKNLRVGRHGLKKRPEEC 660  
  
QY 661 DCHKISYHTQHGRLKHRGSSLHPFRKQKQKWLRLREQKXKLLKRLQNNNDTC 720  
Db 661 DCHKISYHTQHGRLKHRGSSLHPFRKQKQKWLRLREQKXKLLKRLQNNNDTC 720

721 SMPGLTCFTHDNOHWQTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEFATGFE 780  
721 SMPGLTCFTHDNOHWQTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEFATGFE 780  
781 YFDLNTDPQLMANVNTLDRVNLQHLVQLMELRSCKYKQCNPRNMDLGLKGGSYE 840  
781 YFDLNTDPQLMANVNTLDRVNLQHLVQLMELRSCKYKQCNPRNMDLGLKGGSYE 840  
841 QYRQFQRRKWPENKRPSSKSLGQLWEGWEG 870  
841 QYRQFQRRKWPENKRPSSKSLGQLWEGWEG 870

RESULT 2  
AAW79215  
ID AAW79215 standard; Protein; 870 AA.  
XX  
XX AAW79215;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human protein SEQ ID NO 1877.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
XX  
XX WO200157190-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001WO-US04098.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
XX  
XX 27-APR-2000; 2000US-0560875.  
XX  
XX 20-JUN-2000; 2000US-0598075.  
XX  
XX 19-JUL-2000; 2000US-0620325.  
XX  
XX 01-SEP-2000; 2000US-0654936.  
XX  
XX 15-SEP-2000; 2000US-0663561.  
XX  
XX 20-OCT-2000; 2000US-0693325.  
XX  
XX 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
XX WPI; 2001-476283/51.  
XX  
XX N-PSDB; AAK52348.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
XX useful in diagnosis and gene therapy -  
XX  
XX Claim 20; Page 4262-4264; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAW79323-AAW80302) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation.  
XX  
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
XX (AAW80020) are omitted as the relevant pages from the sequence listing  
XX were missing at the time of publication.

[illegible]

CC in diagnostic immuncassays e.g., radioimmunoassay or enzyme linked  
CC immunosorbent assay (ELISA). The present sequence represents a human  
CC secreted protein of the invention.

ABG71036 standard; Protein; 870 AA.  
ABG71836; .  
20-JAN-2003 (first entry)  
human SULF2 protein.  
Human; SULF2; glucosamine-6-sulphatase; cancer; ischaemia; enzyme; tumour; angiogenesis; coronary; carotid; arterial occlusive disease; peripheral arterial disease; atherosclerosis; myointimal hyperplasia; thromboangitis obliterans; thrombotic disorder; vasculitis; heart attack; myocardial infarction; vascular death; inflammation; rheumatoid arthritis; asthma; adult respiratory distress syndrome; sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis; allograft rejection; lymphoma; thrombosis; sulobactase.

CC	distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple
CC	sclerosis, allograft rejection, and spread of lymphomas to cutaneous
CC	sites. The present sequence represents the human SULF2 protein of the
CC	invention.
XX	
XX	
SQ	Sequence 870 AA;
Query Match	100.0%; Score 4750; DB 23; Length 870;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 870; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVELGS 60
DB	1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVELGS 60
QY	61 MQVNMKTRRIMEOGGAHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSFSSWQA 120
DB	61 MQVNMKTRRIMEOGGAHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSFSSWQA 120
QY	121 QHESRTFAVYLNSTGYRTAFFGKYLVNENGSYVPPGKWEVGLLKNRFRNYTLCRNGVK 180
DB	121 QHESRTFAVYLNSTGYRTAFFGKYLVNENGSYVPPGKWEVGLLKNRFRNYTLCRNGVK 180
QY	181 EKHGSDYSKDLTLITNDVSFRTSKMYPHPVLMVISHAAPHGPEDSAPOYSRLFP 240
DB	181 EKHGSDYSKDLTLITNDVSFRTSKMYPHPVLMVISHAAPHGPEDSAPOYSRLFP 240
QY	241 NASQHIPTSYNAPNDKHMIMRTGPMKPIHMEFTNMLQKRLQTLMSVDDSNETIYNN 300
DB	241 NASQHIPTSYNAPNDKHMIMRTGPMKPIHMEFTNMLQKRLQTLMSVDDSNETIYNN 300
QY	301 LVEIGELDNTVIVTADGHYHIGFGLVKGSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360
DB	301 LVEIGELDNTVIVTADGHYHIGFGLVKGSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360
QY	361 LNIIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFLKQORVWRDSFLVERGKLL 420
DB	361 LNIIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFLKQORVWRDSFLVERGKLL 420
QY	421 HKNRNDKVDQAEENFLPKYQVAVKDCQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
DB	421 HKNRNDKVDQAEENFLPKYQVAVKDCQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
QY	481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSTRSV 540
DB	481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSTRSV 540
QY	541 AIEVDGRVYHVLGDAQPRNLTKRHWPFGAPEDDDKGGDFSGTGLPDPYSANPIKVT 600
DB	541 AIEVDGRVYHVLGDAQPRNLTKRHWPFGAPEDDDKGGDFSGTGLPDPYSANPIKVT 600
QY	601 HRCVILENDTVQCDDLYKSLQAKDKHLHDHIEITLQNKIKLNRVRGHLKKRPEEC 660
DB	601 HRCVILENDTVQCDDLYKSLQAKDKHLHDHIEITLQNKIKLNRVRGHLKKRPEEC 660
QY	661 DCHKISYHTQHGKRLKRGSSLHPFRKGLQEKDKVLLRQKRRKKLRLKQLQNNDTC 720
DB	661 DCHKISYHTQHGKRLKRGSSLHPFRKGLQEKDKVLLRQKRRKKLRLKQLQNNDTC 720
QY	721 SMPGLTCTHNDONHQTAPFTWLGPFCACTSANNTYQWRTNETHNLFCFATGFLF 780
DB	721 SMPGLTCTHNDONHQTAPFTWLGPFCACTSANNTYQWRTNETHNLFCFATGFLF 780
QY	781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQMLRSCSKYQKQCNPRTRNMDLGLKDGGSYE 840
DB	781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQMLRSCSKYQKQCNPRTRNMDLGLKDGGSYE 840
QY	841 QYRQFORAKWPEMKRPPSKSLGQLWEGWEG 870
DB	841 QYRQFORAKWPEMKRPPSKSLGQLWEGWEG 870

RESULT 5

ABG63903  
ID ABG63903 standard; Protein; 870 AA.

XX

AC ABG63903;

XX

DT 27-AUG-2002 (first entry)

XX

DE Human albumin fusion protein #578.

XX

KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antiinfertility; antiinflammatory; antitumor;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.

XX

OS Homo sapiens.

OS

Synthetic.

XX

PN WO200177137-A1.

XX

PD 18-OCT-2001.

XX

PF 12-APR-2001; 2001WO-US11988.

XX

PR 12-APR-2000; 2000US-229358P.

PR

25-APR-2000; 2000US-199384P.

PR

21-DEC-2000; 2000US-256931P.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Haseltine WA;

XX

WPI; 2002-010886/01.

XX

New fusion protein for treating disease e.g. diabetes comprises an

XX

Claim 1; Page 892-894; 2102pp; English.

CC

The present invention relates to albumin fusion proteins comprising a  
therapeutic protein X and human albumin (HA), also known as human serum  
albumin, HSA). The proteins are useful for treating a disease or  
disorder that may be modulated by therapeutic protein X. The albumin  
extends the shelf-life of protein X, and may increase its biological  
in vitro/in vivo activity. The protein is useful for treating and  
diagnosing disorders such as cancer, reproductive disorders, digestive  
disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
(e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
(e.g. diabetes), haematopoietic disorders, neural disorders  
(e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
encephalomyelitis, meningitis, schizophrenia), and connective disorders  
(e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
fusion proteins of the invention.

XX

SQ Sequence 870 AA;

Query Match

100.0%; Score 4750; DB 23; Length 870;

Best Local Similarity

100.0%; Pred. No. 0;

Matches 870; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY

1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVELGS 60

DB

1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVELGS 60

QY

61 MQVNMKTRRIMEOGGAHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSFSSWQA 120

DB

61 MQVNMKTRRIMEOGGAHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSFSSWQA 120

QY

121 QHESRTFAVYLNSTGYRTAFFGKYLVNENGSYVPPGKWEVGLLKNRFRNYTLCRNGVK 180

Db 121 QHESRTFAVLNLTGRTAFTGFKYLNEYNGSVPPGKWEVGLLKNSRFYNYTLCRNGVK 180  
Qy 181 EKHGSDYKDYLTDLITNDVSFFETSCKMYPHRPVLWVISHAAPHGPDSPAPQYSLRFP 240  
Db 181 EKHGSDYKDYLTDLITNDVSFFETSCKMYPHRPVLWVISHAAPHGPDSPAPQYSLRFP 240  
Qy 241 NASQHITTSYNAFNPDKHWINRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
Db 241 NASQHITTSYNAFNPDKHWINRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
Qy 301 LVETGELDNITYVTADHGTHYHGOGLVKGSMPEYEDIRVPFVYRGPNVAGCLNPHIV 360  
Db 301 LVETGELDNITYVTADHGTHYHGOGLVKGSMPEYEDIRVPFVYRGPNVAGCLNPHIV 360  
Qy 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERVNFHFKKKKVRVWRDSFLVERGKLL 420  
Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERVNFHFKKKKVRVWRDSFLVERGKLL 420  
Qy 421 HKRDNDKVDQAEENFLPKYQVVDLQRAEYQACEQLGQKQCVEDATGKLLHKCKGP 480  
Db 421 HKRDNDKVDQAEENFLPKYQVVDLQRAEYQACEQLGQKQCVEDATGKLLHKCKGP 480  
Qy 481 MRLGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKLLFKKYYKASVYRGRSRSV 540  
Db 481 MRLGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKLLFKKYYKASVYRGRSRSV 540  
Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHWGAPEDDDXGDFSGTGLPDYSAANPIKVT 600  
Db 541 AIEVDGRVYHVLGDAAPRNLTKRHWGAPEDDDXGDFSGTGLPDYSAANPIKVT 600  
Qy 601 HRCYILENDTVQCDLDLYKSLQAKWDKHLIDHETETLQNKILNREVRGHLKKRPEEC 660  
Db 601 HRCYILENDTVQCDLDLYKSLQAKWDKHLIDHETETLQNKILNREVRGHLKKRPEEC 660  
Qy 661 DCHKLSYTHQHRGLKRGSSLHPFRKQLQEKDQVYMLLREQKQKGLKLLKRLQNNDDTC 720  
Db 661 DCHKLSYTHQHRGLKRGSSLHPFRKQLQEKDQVYMLLREQKQKGLKLLKRLQNNDDTC 720  
Qy 721 SMPGLTCFTHDQHQWQTAFTWTLGPFCACTSANNTYTCMTINETHNPLFCFATGPLE 780  
Db 721 SMPGLTCFTHDQHQWQTAFTWTLGPFCACTSANNTYTCMTINETHNPLFCFATGPLE 780  
Qy 781 YFDLNTDPPQLMNAVNTLDRVLNQLHVMELRSCGYKQCNPTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDPPQLMNAVNTLDRVLNQLHVMELRSCGYKQCNPTRNMDLGLKDGGSYE 840  
Qy 841 QYRQFORRWKPEMKRPSKSLGQLEWEGWEG 870  
Db 841 QYRQFORRWKPEMKRPSKSLGQLEWEGWEG 870

RESULT 6  
AAE01471  
ID AAE01471 standard; Protein; 870 AA.  
XX AAE01471;  
AC AAE01471;  
XX AAE01471;  
DT 17-JUL-2001 (first entry)  
XX Human gene 5 encoded secreted protein HBQ9N39, SEQ ID NO:127.  
DE Human; secreted protein; proliferative disorder; cancer; tumour;  
XX foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
KW endocrine disorder; infection; pregnancy-related disorder;  
KW cell culture; chemotaxis; food additive; gene therapy;  
KW binding partner identification.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..24  
FT /label= Signal\_peptide  
FT Protein 25..870  
FT /note= "Mature human secreted protein"  
FT Misc-difference 537  
FT /label= Unknown  
FT /note= "Encoded by ART"  
XX WC200134626-A1.  
PN 17-MAY-2001.  
XX 01-NOV-2000; 2000WC-US30045.  
XX 05-NOV-1999; 99US-0163581.  
PR 30-JUN-2000; 2000US-0215133.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;  
XX WPI: 2001-308778/32.  
XX N-PSDB; AAD01471.  
XX New nucleic acid molecules encoding 28 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives -  
PS Claim 11; Page 507-510; 562pp; English.  
XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted  
CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.  
CC AAE01514-AAE01544 represent human secreted protein fragments or variants.  
CC The genes and their secreted proteins are useful for preventing,  
CC treating or ameliorating medical conditions, e.g., by protein or gene  
CC therapy. Pathological conditions can be diagnosed by determining the  
CC amount of the new protein in a sample or by determining the presence of  
CC mutations in the new genes. Specific uses are described for each of the  
CC 28 genes, based on the tissues in which they are most highly expressed,  
CC and include developing products for the diagnosis or treatment of  
CC proliferative disorders, cancer, tumours, foetal and developmental  
CC abnormalities, haematopoietic disorders, diseases of the immune system,  
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
CC allergies, neurological disorders (e.g., Alzheimer's disease,  
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
CC cardiovascular disorders, angiogenic disorders, kidney disorders,  
CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
CC disorders, and infections. The proteins can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues, to identify their  
CC cognate ligands or binding partners, and in chemotaxis, and can be used  
CC as a food additive or preservative to modify storage properties.  
CC Antibodies specific for a protein of the invention can be used in  
CC alleviating symptoms associated with the disorders mentioned above, and  
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
CC immunosorbent assay (ELISA). The present sequence represents a human  
CC secreted protein of the invention.  
XX SQ Sequence 870 AA;  
Query Match 99.9%; Score 4746; DB 22; Length 870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MGPPSLVLCILSATVPSLLGSSAFVSHRLKGRFORDRNIRPNILVLTDDQVELGS 60  
Db 1 MGPPSLVLCILSATVPSLLGSSAFVSHRLKGRFORDRNIRPNILVLTDDQVELGS 60

```
QY 61 MQVMKTRIMEQGAHFINAFVTTPMCCPSRSSILTGKYVHNHTYTNNECSPSPWQA 120
DB 61 MQVMKTRIMEQGAHFINAFVTTPMCCPSRSSILTGKYVHNHTYTNNECSPSPWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGWKEWGLKNSRFYNYTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGWKEWGLKNSRFYNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVMSHAAPHGPDSPAPQYSLRFP 240
DB 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVMSHAAPHGPDSPAPQYSLRFP 240
QY 241 NASQHTPSYNAFNPDKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHTPSYNAFNPDKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNITYIVYTADRGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 301 LVETGELDNITYIVYTADRGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRPFLKKNRVRWDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRPFLKKNRVRWDSFLVERGKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWQCVEDATGKLLHKCKGP 480
DB 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWQCVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKYKASVYRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKYKASVYRSRSIRSV 540
QY 541 ALEVDGRVHVGLGDAAPRNLTKEHWPAGPDODDQKGGDFSGTGGLPDYSAANPKVT 600
DB 541 ALEVDGRVHVGLGDAAPRNLTKEHWPAGPDODDQKGGDFSGTGGLPDYSAANPKVT 600
QY 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEITLQNKIKNLRVGRHGLKKRPEEC 660
DB 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEITLQNKIKNLRVGRHGLKKRPEEC 660
QY 661 DCHKLSYTHQKGRUKHGRSSLHPRKGLQEKDWLLREQRKKLKLKRLQNNDC 720
DB 661 DCHKLSYTHQKGRUKHGRSSLHPRKGLQEKDWLLREQRKKLKLKRLQNNDC 720
QY 721 SMPGLTCTFDHNOHWOTAPFTWLTGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
DB 721 SMPGLTCTFDHNOHWOTAPFTWLTGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
QY 781 YPDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELSCSKYKOCNPRTRNMDLGLKDGGSYE 840
DB 781 YPDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELSCSKYKOCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
DB 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
```

## RESULT 7

AB80921

ID AB80921 standard; Protein; 870 AA.

XX AC AB80921;

XX XX

DT DT 08-OCT-2002 (first entry)

XX XX

DE Human sulphatase polypeptide, 22437.

XX XX

KW Human; sulphatase; 22437; cytostatic; vulnerary; neuroprotective;

XX XX

OS Gene therapy; enzyme.

XX XX

OS Homo sapiens.

XX XX

PN WO200252019-A2.

```
XX 04-JUL-2002.
XX 03-OCT-2001; 2001WO-US30856.
XX 21-DEC-2000; 2000US-257082P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Rudolph-Owen LA;
XX WPI; 2002-566677/60.
XX N-PSDB; ABN86352, ABN86353.
XX Identifying agents for modulating (e.g. inhibiting) e.g. tumor
XX establishment, growth or metastases, neuron growth, or wound healing by
XX determining whether a test compound binds with a 22437 polypeptide
XX (human sulfatase) -
XX Claim 19; Fig 1A-F; 143pp; English.
XX The invention relates to identifying a compound useful for modulating at
XX least one phenomenon (e.g. tumour establishment, tumour growth, tumour
XX metastases, epithelial and/or endothelial cell proliferation, neuronal
XX cell growth, wound healing or cerebral injury). The method involves
XX determining whether a test compound binds with a 22437 polypeptide.
XX The identified modulators of 22437 nucleic acid and polypeptide are also
XX useful for treating cancer or wounds (e.g. stroke-related cerebral
XX ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
XX and polypeptide are useful for diagnosing, preventing or treating a
XX subject having cancer or a cellular proliferation and/or differentiation
XX disorder or at risk of developing cancer or a cellular proliferation
XX and/or differentiation disorder. The present sequence represents the
XX human sulphatase polypeptide, 22437.
XX SQ Sequence 870 AA;
```

```
Query Match 99.9%; Score 4746; DB 23; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPPSLVLCCLLSATVFLSLLGSSAFLSHRLKGRFQDRRIRNPNIILVLTDDQDVELGS 60
DB 1 MGPPSLVLCCLLSATVFLSLLGSSAFLSHRLKGRFQDRRIRNPNIILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEQGAHFINAFVTTPMCCPSRSSILTGKYVHNHTYTNNECSPSPWQA 120
DB 61 MQVMKTRIMEQGAHFINAFVTTPMCCPSRSSILTGKYVHNHTYTNNECSPSPWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGWKEWGLKNSRFYNYTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGWKEWGLKNSRFYNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVMSHAAPHGPDSPAPQYSLRFP 240
DB 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVMSHAAPHGPDSPAPQYSLRFP 240
QY 241 NASQHTPSYNAFNPDKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHTPSYNAFNPDKHWMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNITYIVYTADRGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 301 LVETGELDNITYIVYTADRGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRPFLKKNRVRWDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRPFLKKNRVRWDSFLVERGKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWQCVEDATGKLLHKCKGP 480
DB 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWQCVEDATGKLLHKCKGP 480
```



QY 481 MRLGSRALSNLVKPYQYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540  
DB 481 MRLGSRALSNLVKPYQYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540  
QY 541 AIEVDGRVYHVLGDAAPRNLTQRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
DB 541 AIEVDGRVYHVLGDAAPRNLTQRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
QY 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKDKVWLLRQKRRKKLRLKRLQNDTC 720  
DB 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKDKVWLLRQKRRKKLRLKRLQNDTC 720  
QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFATGFL 780  
DB 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFATGFL 780  
QY 781 YFDLNTDPQLMNAVNTLDRDVLNQLHVQLMELRCKGYKQCNPRTRNMDLGLKDGGSYE 840  
DB 781 YFDLNTDPQLMNAVNTLDRDVLNQLHVQLMELRCKGYKQCNPRTRNMDLGLKDGGSYE 840  
QY 841 QYRQFORRWPEKPPSSKSLGQWEGWEG 870  
DB 841 QYRQFORRWPEKPPSSKSLGQWEGWEG 870

## RESULT 8

ID ABG63902  
XX ABG63902 standard; Protein; 870 AA.  
AC ABG63902;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Human albumin fusion protein #577.  
XX  
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antiinfectivity; antiinflammatory; antitumor;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neotropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200177137-A1.  
XX  
PD 18-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US11988.  
XX  
PR 12-APR-2000; 2000US-229358P.  
PR 25-APR-2000; 2000US-199384P.  
PR 21-DEC-2000; 2000US-256931P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
PI WPI; 2002-010886/01.  
XX  
DR  
XX  
PT New fusion protein for treating disease e.g. diabetes comprises an  
PT albumin fused to a therapeutic protein -  
XX  
PS Claim 1; Page 889-891; 2102pp; English.  
XX  
CC The present invention relates to albumin fusion proteins comprising a

therapeutic protein X and human albumin (HA, also known as human serum albumin, HSA). The proteins are useful for treating a disease or disorder that may be modulated by therapeutic protein X. The albumin extends the shelf-life of protein X, and may increase its biological in vitro/in vivo activity. The protein is useful for treating and diagnosing disorders such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes), haematopoietic disorders, neural disorders (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis, schizophrenia), and connective disorders (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin fusion proteins of the invention.

Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 23; Length 870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60  
DB 1 MGPPSLVLCCLLSATVPSLLGGSSAFLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60  
QY 61 MQVMKTRIMEQGGAHFINAVFTTMCPCSSSILTKGYVHNHTYTNNECSSPSQQA 120  
DB 61 MQVMKTRIMEQGGAHFINAVFTTMCPCSSSILTKGYVHNHTYTNNECSSPSQQA 120  
QY 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYNGSVVFPQKWEVGLLKNSRFPYNTLCRNGYK 180  
DB 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYNGSVVFPQKWEVGLLKNSRFPYNTLCRNGYK 180  
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQYPHRVLWVISHAAPHGSDSAPQYSLRFP 240  
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQYPHRVLWVISHAAPHGSDSAPQYSLRFP 240  
QY 241 NASOHTTPSYNAPNPDKEWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
DB 241 NASOHTTPSYNAPNPDKEWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGBELNTYIVYADHYHIGQGLVKGKSMPEYFDIRVPFYVGRGNVEAGCLNPHIV 360  
DB 301 LVETGBELNTYIVYADHYHIGQGLVKGKSMPEYFDIRVPFYVGRGNVEAGCLNPHIV 360  
QY 361 LNIDLAPTLDIAGLDIPADMDGKSLILKLLDTERPVRFLKKQKVRWDSFLVERGKLL 420  
DB 361 LNIDLAPTLDIAGLDIPADMDGKSLILKLLDTERPVRFLKKQKVRWDSFLVERGKLL 420  
QY 421 HKRDNDKVDQAEENFLPKYQYRVKDLQRAEYQTAECQLGQKQCVEDATGKILKHKCKGP 480  
DB 421 HKRDNDKVDQAEENFLPKYQYRVKDLQRAEYQTAECQLGQKQCVEDATGKILKHKCKGP 480  
QY 481 MRLGSRALSNLVKPYQYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540  
DB 481 MRLGSRALSNLVKPYQYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540  
QY 541 AIEVDGRVYHVLGDAAPRNLTQRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
DB 541 AIEVDGRVYHVLGDAAPRNLTQRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
QY 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKDKVWLLRQKRRKKLRLKRLQNDTC 720  
DB 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKDKVWLLRQKRRKKLRLKRLQNDTC 720  
QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFATGFL 780  
DB 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFATGFL 780  
QY 781 YFDLNTDPQLMNAVNTLDRDVLNQLHVQLMELRCKGYKQCNPRTRNMDLGLKDGGSYE 840

Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVLMLRSCKYKQCNPRTRNMDLGLKGGGSYE 840  
QY 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870  
Db 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870

RESULT 9  
ABB90378  
ID ABB90378 standard; Protein; 870 AA.  
XX  
AC ABB90378;  
DT 24-MAY-2002 (first entry)  
DE Human polypeptide SEQ ID NO 2754.  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW aniallergic; hepatotropic; antidiabetic; antinflammatory; antiulcer;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.  
XX WO200190304-A2.  
XX 29-NOV-2001.  
XX 18-MAY-2001; 2001WO-US16450.  
XX 19-MAY-2000; 2000US-205515P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Birse CE, Rosen CA;  
XX WPI; 2002-122018/16.  
XX N-PSDB; ABL90787.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -  
Claim 11; SEQ ID NO 2754; 2081pp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune disease, multiple sclerosis, rheumatoid arthritis, diabetes mellitus, Crohn's disease, hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 870 AA;  
Query Match 99.9%; Score 4746; DB 23; Length 870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVCLLSATVFSLLGGSSAFLSHERLKGRRFDRDRNIRPNIIILVLTDDQDVELGS 60  
Db 1 MGPPSLVCLLSATVFSLLGGSSAFLSHERLKGRRFDRDRNIRPNIIILVLTDDQDVELGS 60  
QY 61 MQVMKTRIMEQGAHFINAFVTTMCCPSSSSILTKGYVHNHTYNNENCSPPSQA 120  
Db 61 MQVMKTRIMEQGAHFINAFVTTMCCPSSSSILTKGYVHNHTYNNENCSPPSQA 120  
QY 121 QHESRTFAVILNSTGYRTAFGKYLNEVNGSVPPGKWEKGLLNKSNRYNNTLTCRNGVK 180  
Db 121 QHESRTFAVILNSTGYRTAFGKYLNEVNGSVPPGKWEKGLLNKSNRYNNTLTCRNGVK 180  
QY 181 EKXGSDYSKOYLTDLITNDSVSFFRTSKMTPHRPVLVMSHAAPHPGDESAPOYSRLFF 240  
Db 181 EKXGSDYSKOYLTDLITNDSVSFFRTSKMTPHRPVLVMSHAAPHPGDESAPOYSRLFF 240  
QY 241 NASOHITPSYVAPNDKHMRYTGPMKPIHMETNMLORKELOTLMSVDDSMETIYNN 300  
Db 241 NASOHITPSYVAPNDKHMRYTGPMKPIHMETNMLORKELOTLMSVDDSMETIYNN 300  
QY 301 LVETGELONTYIVYTADHGYHIGQFLVKGKSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
Db 301 LVETGELONTYIVYTADHGYHIGQFLVKGKSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
QY 361 LNIDLAPTILDIAGLDIPADMDGKILKLDTERPVNRFHLKKGRVWRDPLVERGKLL 420  
Db 361 LNIDLAPTILDIAGLDIPADMDGKILKLDTERPVNRFHLKKGRVWRDPLVERGKLL 420  
QY 421 HKRDNKVDAAQENFLPKYQVRKDLCCRAEYQTAQCEQLGQKWCQVEDATGKLKHKCKGP 480  
Db 421 HKRDNKVDAAQENFLPKYQVRKDLCCRAEYQTAQCEQLGQKWCQVEDATGKLKHKCKGP 480  
QY 481 MRLGGSRALSNLVPKYQGSEACTCDSGYKLSLAGRRKKLKKKYKASYVRSRSISV 540  
Db 481 MRLGGSRALSNLVPKYQGSEACTCDSGYKLSLAGRRKKLKKKYKASYVRSRSISV 540  
QY 541 ALEVDGRVTHVGLGDAQPRNLTKHWPAGPEDDDDXGDFSGTGGLPDYSAANPIKVT 600  
Db 541 ALEVDGRVTHVGLGDAQPRNLTKHWPAGPEDDDDXGDFSGTGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCOLDLYKSLQAWKDHKLHIDHETLQNKIKNLREVRGHLKKRPREEC 660  
Db 601 HRCYILENDTVQCOLDLYKSLQAWKDHKLHIDHETLQNKIKNLREVRGHLKKRPREEC 660  
QY 661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDWLLREQRKKKLLKRLQNNDTC 720  
Db 661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDWLLREQRKKKLLKRLQNNDTC 720  
QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFL 780  
Db 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFL 780  
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVLMLRSCKYKQCNPRTRNMDLGLKGGGSYE 840  
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVLMLRSCKYKQCNPRTRNMDLGLKGGGSYE 840  
QY 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870  
Db 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870  
RESULT 10  
ABB80922  
ID ABB80922 standard; Protein; 885 AA.  
XX  
AC ABB80922;  
DT 08-OCT-2002 (first entry)  
DE Human KIAA1427 polypeptide.  
XX  
KW Human; sulphatase; 22437; cytostatic; vulnery; neuroprotective;  
KW gene therapy; KIAA1427.

XX OS Homo sapiens.  
XX PN WO200252019-A2.  
XX PD 04-JUL-2002.  
XX PF 03-OCT-2001; 2001WO-US30856.  
XX PR 21-DEC-2000; 2000US-257082P.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX FI Glucksmann MA, Rudolph-Owen LA;  
XX DR WPI; 2002-566677/60.  
XX DR N-PSDB; AEN86354.  
XX PT Identifying agents for modulating (e.g. inhibiting) e.g. tumor  
XX PT establishment, growth or metastases, neuron growth, or wound healing by  
XX PT determining whether a test compound binds with a 22437 polypeptide  
XX PS (human sulfatase) -  
XX PS Disclosure; Fig 4A-D; 143pp; English.  
XX CC The invention relates to identifying a compound useful for modulating at  
XX CC least one phenomenon (e.g. tumor establishment, tumor growth, tumor  
XX CC metastases, epithelial and/or endothelial cell proliferation, neuronal  
XX CC cell growth, wound healing or cerebral injury). The method involves  
XX CC determining whether a test compound binds with a 22437 polypeptide.  
XX CC The identified modulators of 22437 nucleic acid and polypeptide are also  
XX CC useful for treating cancer or wounds (e.g. stroke-related cerebral  
XX CC ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid  
XX CC and polypeptide are useful for diagnosing, preventing or treating a  
XX CC subject having cancer or a cellular proliferation and/or differentiation  
XX CC disorder or at risk of developing cancer or a cellular proliferation  
XX CC and/or differentiation disorder. The present sequence represents a human  
XX CC KIAA1427 polypeptide used in alignment studies with the human sulphatase  
XX CC polypeptide, 22437.  
XX SQ Sequence 885 AA;

Query Match 99.9%; Score 4746; DB 23; Length 885;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRFQDRDRNIRPNIILVLTDDQVELGS 60  
DB 16 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRFQDRDRNIRPNIILVLTDDQVELGS 75  
QY 61 MOVNKTERRIMEOGGAHFINAFVTPMCCPSRSSILTKGVHNTVYNNENSSPSWQA 120  
DB 76 MOVNKTERRIMEOGGTHFINAFVTPMCCPSRSSILTKGVHNTVYNNENSSPSWQA 135  
QY 121 QHESRTFAVLNLTGYRTAFKGLYNEVNGSYVPPGKWEVGLLNKSRFYNYTLCRNGVK 180  
DB 136 QHESRTFAVLNLTGYRTAFKGLYNEVNGSYVPPGKWEVGLLNKSRFYNYTLCRNGVK 195  
QY 181 EKHGSDYSKDLVLTDLITNDSVSPFRTSKMYPHPVPLVWISHAAPHGPDSPAQYSLRFP 240  
DB 196 EKHGSDYSKDLVLTDLITNDSVSPFRTSKMYPHPVPLVWISHAAPHGPDSPAQYSLRFP 255  
QY 241 NASQHTPSYNAFPDCKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
DB 256 NASQHTPSYNAFPDCKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 315  
QY 301 LVETGELDNTVIVYADGHYHIGFGLVKGSMEYEDIRVFFYVRGPNVEAGCLAPHIV 360  
DB 316 LVETGELDNTVIVYADGHYHIGFGLVKGSMEYEDIRVFFYVRGPNVEAGCLAPHIV 375  
QY 361 LNIIDLAPITLIDAGLIDIPADMDGKSILKLDTERPVNRFHLKKQVRWRDSFLVERGKLL 420  
DB 376 LNIIDLAPITLIDAGLIDIPADMDGKSILKLDTERPVNRFHLKKQVRWRDSFLVERGKLL 435

QY 421 HKRDNDKYDAQENFLPKYQVRKDLCPRAEYQTACEQLGQKWCQVEDATGKLKHKCKGP 480  
DB 436 HKRDNDKYDAQENFLPKYQVRKDLCPRAEYQTACEQLGQKWCQVEDATGKLKHKCKGP 495  
QY 481 MRLGGSRALSNLVPKYQGSGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSRSV 540  
DB 496 MRLGGSRALSNLVPKYQGSGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSRSV 555  
QY 541 ALEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODKGGDFSGTGGLPDYSAANPIKVT 600  
DB 556 ALEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODKGGDFSGTGGLPDYSAANPIKVT 615  
QY 601 HRCYILENDTVQCCLDLKYSLOAWKHKLHIDHIEITLQNKIKNLREYRHLKKRPEEC 660  
DB 616 HRCYILENDTVQCCLDLKYSLOAWKHKLHIDHIEITLQNKIKNLREYRHLKKRPEEC 675  
QY 661 DCHKISYHTQHKGLKRGSSIHPRKGLQEKVKVLLREQKRRKKLKLKRLQNNDTG 720  
DB 676 DCHKISYHTQHKGLKRGSSIHPRKGLQEKVKVLLREQKRRKKLKLKRLQNNDTG 735  
QY 721 SNPGLTCETHDNQHWQAPFWTLGPFCACTSANNTYWCMTINETINFLFCEPATGFLE 780  
DB 736 SNPGLTCETHDNQHWQAPFWTLGPFCACTSANNTYWCMTINETINFLFCEPATGFLE 795  
QY 781 YFDLNTDTPYQLMNAVNTLDRVLNQLHVLQVLMELRSCKGYKOCNPRTRNMDLGLKGGSYE 840  
DB 796 YFDLNTDTPYQLMNAVNTLDRVLNQLHVLQVLMELRSCKGYKOCNPRTRNMDLGLKGGSYE 855  
QY 841 QYRQFORAKWPEMKRPPSSKSLGQLWEGWEG 870  
DB 856 QYRQFORAKWPEMKRPPSSKSLGQLWEGWEG 885

RESULT 11  
AAAY66648  
ID AAAY66648 standard; protein; 867 AA.  
XX  
AC AAAY66648;  
XX  
DT 05-APR-2000 (first entry)  
XX  
DE Membrane-bound protein PRO1120.  
XX  
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
KW pharmaceutical; receptor immunoadhesin; gene mapping.  
XX  
OS Homo sapiens.  
XX  
PN WO9963088-A2.  
XX  
PD  
XX  
XX 09-DEC-1999.  
XX  
PF 02-JUN-1999; 99WO-US12252.  
XX  
XX 02-JUN-1998; 98US-0087607.  
PR 02-JUN-1998; 98US-0087609.  
PR 02-JUN-1998; 98US-0087759.  
PR 03-JUN-1998; 98US-0087827.  
PR 04-JUN-1998; 98US-0088021.  
PR 04-JUN-1998; 98US-0088025.  
PR 04-JUN-1998; 98US-0088028.  
PR 04-JUN-1998; 98US-0088029.  
PR 04-JUN-1998; 98US-0088030.  
PR 04-JUN-1998; 98US-0088033.  
PR 04-JUN-1998; 98US-0088326.  
PR 05-JUN-1998; 98US-0088167.  
PR 05-JUN-1998; 98US-0088202.  
PR 05-JUN-1998; 98US-0088212.  
PR 05-JUN-1998; 98US-0088217.  
PR 09-JUN-1998; 98US-0088655.  
PR 10-JUN-1998; 98US-0088722.  
PR 10-JUN-1998; 98US-0088730.

PR 10-JUN-1998; 98US-0088734.  
 PR 10-JUN-1998; 98US-0088738.  
 PR 10-JUN-1998; 98US-0088740.  
 PR 10-JUN-1998; 98US-0088741.  
 PR 10-JUN-1998; 98US-0088742.  
 PR 10-JUN-1998; 98US-0088810.  
 PR 10-JUN-1998; 98US-0088811.  
 PR 10-JUN-1998; 98US-0088824.  
 PR 10-JUN-1998; 98US-0088825.  
 PR 10-JUN-1998; 98US-0088826.  
 PR 11-JUN-1998; 98US-0088858.  
 PR 11-JUN-1998; 98US-0088861.  
 PR 11-JUN-1998; 98US-0088863.  
 PR 11-JUN-1998; 98US-0088876.  
 PR 12-JUN-1998; 98US-0089090.  
 PR 12-JUN-1998; 98US-0089105.  
 PR 16-JUN-1998; 98US-0089440.  
 PR 16-JUN-1998; 98US-0089512.  
 PR 16-JUN-1998; 98US-0089514.  
 PR 17-JUN-1998; 98US-0089532.  
 PR 17-JUN-1998; 98US-0089538.  
 PR 17-JUN-1998; 98US-0089598.  
 PR 17-JUN-1998; 98US-0089599.  
 PR 17-JUN-1998; 98US-0089600.  
 PR 17-JUN-1998; 98US-0089653.  
 PR 18-JUN-1998; 98US-0089801.  
 PR 18-JUN-1998; 98US-0089907.  
 PR 18-JUN-1998; 98US-0089908.  
 PR 19-JUN-1998; 98US-0089947.  
 PR 19-JUN-1998; 98US-0089948.  
 PR 19-JUN-1998; 98US-0089952.  
 PR 22-JUN-1998; 98US-0090246.  
 PR 22-JUN-1998; 98US-0090252.  
 PR 23-JUN-1998; 98US-0090254.  
 PR 23-JUN-1998; 98US-0090349.  
 PR 23-JUN-1998; 98US-0090355.  
 PR 24-JUN-1998; 98US-0090429.  
 PR 24-JUN-1998; 98US-0090431.  
 PR 24-JUN-1998; 98US-0090435.  
 PR 24-JUN-1998; 98US-0090444.  
 PR 24-JUN-1998; 98US-0090445.  
 PR 24-JUN-1998; 98US-0090451.  
 PR 24-JUN-1998; 98US-0090472.  
 PR 24-JUN-1998; 98US-0090535.  
 PR 24-JUN-1998; 98US-0090538.  
 PR 24-JUN-1998; 98US-0090540.  
 PR 24-JUN-1998; 98US-0090557.  
 PR 25-JUN-1998; 98US-0090676.  
 PR 25-JUN-1998; 98US-0090678.  
 PR 25-JUN-1998; 98US-0090688.  
 PR 25-JUN-1998; 98US-0090690.  
 PR 25-JUN-1998; 98US-0090691.  
 PR 25-JUN-1998; 98US-0090694.  
 PR 25-JUN-1998; 98US-0090695.  
 PR 25-JUN-1998; 98US-0090696.  
 PR 26-JUN-1998; 98US-0090862.  
 PR 26-JUN-1998; 98US-0090863.  
 PR 01-JUL-1998; 98US-0091358.  
 PR 01-JUL-1998; 98US-0091360.  
 PR 01-JUL-1998; 98US-0091544.  
 PR 02-JUL-1998; 98US-0091478.  
 PR 02-JUL-1998; 98US-0091486.  
 PR 02-JUL-1998; 98US-0091519.  
 PR 02-JUL-1998; 98US-0091626.  
 PR 02-JUL-1998; 98US-0091628.  
 PR 02-JUL-1998; 98US-0091633.  
 PR 02-JUL-1998; 98US-0091646.  
 PR 02-JUL-1998; 98US-0091673.  
 PR 07-JUL-1998; 98US-0091978.  
 PR 07-JUL-1998; 98US-0091982.  
 PR 09-JUL-1998; 98US-0092182.  
 PR 10-JUL-1998; 98US-0092472.  
 PR 20-JUL-1998; 98US-0093339.

PR 30-JUL-1998; 98US-0094651.  
 PR 04-AUG-1998; 98US-0095282.  
 PR 04-AUG-1998; 98US-0095285.  
 PR 04-AUG-1998; 98US-0095301.  
 PR 04-AUG-1998; 98US-0095302.  
 PR 04-AUG-1998; 98US-0095318.  
 PR 04-AUG-1998; 98US-0095321.  
 PR 04-AUG-1998; 98US-0095325.  
 PR 10-AUG-1998; 98US-0095916.  
 PR 10-AUG-1998; 98US-0095929.  
 PR 10-AUG-1998; 98US-0096012.  
 PR 11-AUG-1998; 98US-0096143.  
 PR 11-AUG-1998; 98US-0096146.  
 PR 12-AUG-1998; 98US-0096329.  
 PR 17-AUG-1998; 98US-0096757.  
 PR 17-AUG-1998; 98US-0096766.  
 PR 17-AUG-1998; 98US-0096768.  
 PR 17-AUG-1998; 98US-0096773.  
 PR 17-AUG-1998; 98US-0096791.  
 PR 17-AUG-1998; 98US-0096867.  
 PR 17-AUG-1998; 98US-0096891.  
 PR 17-AUG-1998; 98US-0096894.  
 PR 17-AUG-1998; 98US-0096895.  
 PR 17-AUG-1998; 98US-0096897.  
 PR 18-AUG-1998; 98US-0096949.  
 PR 18-AUG-1998; 98US-0096950.  
 PR 18-AUG-1998; 98US-0096959.  
 PR 18-AUG-1998; 98US-0096960.  
 PR 18-AUG-1998; 98US-0097022.  
 PR 20-AUG-1998; 98US-0097141.  
 PR 20-AUG-1998; 98US-0097218.  
 PR 24-AUG-1998; 98US-0097661.  
 PR 26-AUG-1998; 98US-0097951.  
 PR 26-AUG-1998; 98US-0097952.  
 PR 26-AUG-1998; 98US-0097954.  
 PR 26-AUG-1998; 98US-0097955.  
 PR 26-AUG-1998; 98US-0097971.  
 PR 26-AUG-1998; 98US-0097974.  
 PR 26-AUG-1998; 98US-0097978.  
 PR 26-AUG-1998; 98US-0097979.  
 PR 26-AUG-1998; 98US-0097986.  
 PR 26-AUG-1998; 98US-0098014.  
 PR 31-AUG-1998; 98US-0098525.  
 PR 16-SEP-1998; 98US-0100634.  
 PR 12-JAN-1999; 98US-0115565.  
 XX  
 PA (GETH ) GENENTECH INC.

XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
 PI Wood WI, Yuan J;  
 XX WPI: 2000-072883/06.  
 DR N-PSDB; AAZ64969.  
 DR

Membrane-bound proteins and related nucleotide sequences -  
 claim 12; Fig 47; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

XX

SQ	Sequence	867 AA;
QY	Query Match	99.4%; Score 4723.5; DB 21; Length 867;
QY	Best Local Similarity	99.7%; Pred. No. 0;
QY	Matches 867; Conservative	0; Mismatches 0; Indels 3; Gaps 1;
QY	1	MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFDRNRNIRNIIILVLTDDQDVELGS 60
QY	1	MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFDRNRNIRNIIILVLTDDQDVELGS 60
QY	61	MOVNNKTRIMEOGGAHTINAFVTTMCCPSRSSLTGKYVHNNTYNNENCSPPSQA 120
QY	61	MOVNNKTRIMEOGGAHTINAFVTTMCCPSRSSLTGKYVHNNTYNNENCSPPSQA 120
QY	121	QHSRTFAVLNSTGYRTAFKGYLNEYNGSVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
QY	121	QHSRTFAVLNSTGYRTAFKGYLNEYNGSVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
QY	181	EKHGSDYKDYLTDLITDYSVSFFRTSKMYPHRPVLVWISHAAPHGPDAPQYSRLFP 240
QY	181	EKHGSDYKDYLTDLITDYSVSFFRTSKMYPHRPVLVWISHAAPHGPDAPQYSRLFP 240
QY	241	NASQHTSYNAPNPDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY	241	NASQHTSYNAPNPDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY	301	LVEGTGLDNTIYVYADHGYHIGQGLVKGKSMPEYFDIRVFPVYRGPNVBAAGCLNPHIV 360
QY	301	LVEGTGLDNTIYVYADHGYHIGQGLVKGKSMPEYFDIRVFPVYRGPNVBAAGCLNPHIV 360
QY	361	LNIDLAFTILDAGLDIPADMDGKILKLDTERPVNRFHLKXKRWVWDSFLVERGKLL 420
QY	361	LNIDLAFTILDAGLDIPADMDGKILKLDTERPVNRFHLKXKRWVWDSFLVERGKLL 420
QY	421	HKRDNDKVDAAEENFLPKYQYKDIQCPAEYQTAQCEQLGQKQWQVEDATGKLKHKCKGP 480
QY	421	HKRDNDKVDAAEENFLPKYQYKDIQCPAEYQTAQCEQLGQKQWQVEDATGKLKHKCKGP 480
QY	481	MLGGSRLSNLVPKYGGSGSACTCDSDGYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
QY	481	MLGGSRLSNLVPKYGGSGSACTCDSDGYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
QY	541	AIEVDGRVYHVLGDAAPRNITKRWFGAPEDQDDKDGDFSGTGGLPDYSAANPKVIT 600
QY	541	AIEVDGRVYHVLGDAAPRNITKRWFGAPEDQDDKDGDFSGTGGLPDYSAANPKVIT 600
QY	601	HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHGLKKRPEEC 660
QY	601	HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHGLKKRPEEC 660
QY	661	DCHKISYHTQHKGRGLKGRGSSLHPRKGLQEKDKVWLLREOKRKKLKLRLQNNDTC 720
QY	661	DCHKISYHTQHKGRGLKGRGSSLHPRKGLQEKDKVWLLREOKRKKLKLRLQNNDTC 720
QY	721	SMPGLTCTFHDNQHWQTAFTWTLGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780
QY	721	SMPGLTCTFHDNQHWQTAFTWTLGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780
QY	781	YFDLNTDPQLMNAVNTLDRDLNGLHVLMLRSCKYKQCNPTRNMDLGLXDGGSYE 840
QY	781	YFDLNTDPQLMNAVNTLDRDLNGLHVLMLRSCKYKQCNPTRNMDLGLXDGGSYE 840
QY	841	QYRQFORRWKPEKRPSSLSQLWEGWEG 870
QY	838	QYRQFORRWKPEKRPSSLSQLWEGWEG 867

RESULT 12

AAU12181

ID AAU12181 standard; Protein; 867 AA.

XX

AC

AAU12181;

XX

DT	24-OCT-2001	(first entry)
XX	Human PRO1120 polypeptide sequence.	
XX	Human secretory and transmembrane; PRO; mammalian; cancer; lung;	
XX	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;	
XX	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;	
XX	adipocyte; A-peptide; factor VIIA; gene therapy.	
XX	Homo sapiens.	
XX	WO200140466-A2.	
XX	07-JUN-2001.	
XX	01-DEC-2000; 2000WO-US32678.	
XX	01-DEC-1999; 99WO-US28301.	
XX	01-DEC-1999; 99WO-US28634.	
XX	02-DEC-1999; 99WO-US28551.	
XX	02-DEC-1999; 99WO-US28564.	
XX	02-DEC-1999; 99WO-US28565.	
XX	09-DEC-1999; 99US-0170262.	
XX	16-DEC-1999; 99WO-US30095.	
XX	20-DEC-1999; 99WO-US30911.	
XX	20-DEC-1999; 99WO-US30999.	
XX	30-DEC-1999; 99WO-US31243.	
XX	06-JAN-2000; 2000WO-US00277.	
XX	06-JAN-2000; 2000WO-US00376.	
XX	11-FEB-2000; 2000WO-US03565.	
XX	18-FEB-2000; 2000WO-US04341.	
XX	18-FEB-2000; 2000WO-US04342.	
XX	22-FEB-2000; 2000WO-US04414.	
XX	24-FEB-2000; 2000WO-US04914.	
XX	24-FEB-2000; 2000WO-US05004.	
XX	01-MAR-2000; 2000WO-US05601.	
XX	20-MAR-2000; 2000WO-US07377.	
XX	21-MAR-2000; 2000WO-US07532.	
XX	30-MAR-2000; 2000WO-US08439.	
XX	17-MAY-2000; 2000WO-US13705.	
XX	22-MAY-2000; 2000WO-US14042.	
XX	30-MAY-2000; 2000WO-US14941.	
XX	02-JUN-2000; 2000WO-US15264.	
XX	10-NOV-2000; 2000WO-US30873.	
XX	(GETH ) GENENTECH INC.	
XX	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;	
XX	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;	
XX	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;	
XX	WPI; 2001-408281/43.	
XX	N-PSDB; AAS21253.	
XX	Isolated, secretory and transmembrane PRO polypeptide used to detect	
XX	other PRO polypeptides, link bioactive molecules to cells expressing	
XX	PRO polypeptides, and detect the presence of mammalian tumours e.g.	
XX	lung, breast, prostate, cervical	
XX	Claim 12; Fig 20; 813pp; English.	
XX	AAU12172-AAU12446 represent novel human secretory and transmembrane	
XX	PRO polypeptides. The PRO polypeptides are useful to detect other	
XX	PRO polypeptides, to link bioactive molecules to cells expressing	
XX	PRO polypeptides, to modulate biological activities of cells expressing	
XX	PRO polypeptides, and to detect the presence of mammalian lung, colon,	
XX	breast, prostate, rectal, cervical or liver tumours by comparing PRO	
XX	polypeptide expression in a cell sample to that in a control sample.	
XX	Some of the 275 sequences are also useful to stimulate the release of	
XX	tumour necrosis factor-alpha (TNF-alpha) from human blood, the	
XX	proliferation or differentiation of chondrocytes, the proliferation or	
XX	gene expression in pericyte cells, the release of proteoglycans from	
XX	cartilage, the proliferation of inner ear utricular supporting cells or	

CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX  
SQ Sequence 867 AA;

Query Match		99.4%;	Score 4723.5;	DB 22;	Length 867;
Best Local Similarity		99.7%;	Pred. No. 0;		
Matches 867;		Conservative 0;	Mismatches 0;	Indels 3;	Gaps 1;
QY	1	MGPPSLVLCISATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVBLGS	60		
DB	1	MGPPSLVLCISATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQVBLGS	60		
QY	61	MQVMKTRIRIMEOGGAHFINAVFTPMCCPSRSSILTKGYVHNNTYNNENCSSPSWQA	120		
DB	61	MQVMKTRIRIMEOGGAHFINAVFTPMCCPSRSSILTKGYVHNNTYNNENCSSPSWQA	120		
QY	121	QHSRTFAVILNSTGYRTAFPGKVLNEYNGSVYPPGKWEVGLLKNSRFNYTLCRNGVK	180		
DB	121	QHSRTFAVILNSTGYRTAFPGKVLNEYNGSVYPPGKWEVGLLKNSRFNYTLCRNGVK	180		
QY	181	EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHPVLMVISHAAPHGPDSPAPQSYSLFP	240		
DB	181	EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHPVLMVISHAAPHGPDSPAPQSYSLFP	240		
QY	241	NASQHITPSYNAPNPDKHMVITGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN	300		
DB	241	NASQHITPSYNAPNPDKHMVITGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN	300		
QY	301	LVEGTCLDNTIIVYTADHGYHIGFGLVKGSMZYEFDIRVPYVGRPNVEAGCLNPHIV	360		
DB	301	LVEGTCLDNTIIVYTADHGYHIGFGLVKGSMZYEFDIRVPYVGRPNVEAGCLNPHIV	360		
QY	361	LNIDLAPILDIADMDGKSLKLDTEPVRNREHLKKRVRWDSFLVERKLL	420		
DB	361	LNIDLAPILDIADMDGKSLKLDTEPVRNREHLKKRVRWDSFLVERKLL	420		
QY	421	HKRDNKVDAAEENFLPKYQRYKOLQRAEYQTAQCEQLGQKWCQVEDATGKLHKCKGCP	480		
DB	421	HKRDNKVDAAEENFLPKYQRYKOLQRAEYQTAQCEQLGQKWCQVEDATGKLHKCKGCP	480		
QY	481	WELGGSRLSNLVPKYGGSGEACTCDSDGYKLSLAGEKELFKKKYKASYRVSRSRSV	540		
DB	481	WELGGSRLSNLVPKYGGSGEACTCDSDGYKLSLAGEKELFKKKYKASYRVSRSRSV	540		
QY	541	ALVVDGRVYHVLGDAAPRNLTKEHWPFGAPEDQDDGDFGSGTGLPDYSAANPIKVT	600		
DB	541	ALVVDGRVYHVLGDAAPRNLTKEHWPFGAPEDQDDGDFGSGTGLPDYSAANPIKVT	600		
QY	601	HRCYILENTVQCDDLDLYKSLQANKDKHLHDHIEIETLQNKIKNLRVGRHLKGRPEEC	660		
DB	601	HRCYILENTVQCDDLDLYKSLQANKDKHLHDHIEIETLQNKIKNLRVGRHLKGRPEEC	660		
QY	661	DCHKLSYTHQKGRUKHKGSSLHPPKGLQEKDKVLLREOKRKKLKLKRLQNNNTC	720		
DB	661	DCHKLSYTHQKGRUKHKGSSLHPPKGLQEKDKVLLREOKRKKLKLKRLQNNNTC	720		
QY	721	SNPGLTCTHNDQHWQTAFTWTLGSPFCACTSANNTYWCMTINETNFIKCEPATGFLE	780		
DB	721	SNPGLTCTHNDQHWQTAFTWTLGSPFCACTSANNTYWCMTINETNFIKCEPATGFLE	780		
QY	781	YFDLNTDYPQLMNAVNTLDRVLNQLHVLMELSCKYKOCNPRTRNMDLGLKDGGSYE	840		
DB	781	YFDLNTDYPQLMNAVNTLDRVLNQLHVLMELSCKYKOCNPRTRNMDLGLKDGGSYE	840		
QY	841	QYRQFQRRKWPMPKSPSSKSLQGLWEGWEG	870		

Db 838 QYRQFQRRKWPMPKSPSSKSLQGLWEGWEG 867

RESULT 13

AAB65171

ID AAB65171 standard; Protein; 867 AA.

XX AC AAB65171;

XX DT 02-APR-2001 (first entry)

XX DE Human PRO1120 (UNQ559) protein sequence SEQ ID NO:84.

XX KW Human; secreted and transmembrane protein; PRO; cyrostatic;  
cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
diagnostic assay.

XX OS Homo sapiens.

XX PN WO200073454-A1.

XX PD 07-DEC-2000.

XX PF 30-MAR-2000; 2000WO-US08439.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 23-JUN-1999; 99US-0141037.

XX PR 07-JUL-1999; 99US-0143048.

XX PR 20-JUL-1999; 99US-0144758.

XX PR 26-JUL-1999; 99US-0145698.

XX PR 28-JUL-1999; 99US-0146222.

XX PR 17-AUG-1999; 99US-0149396.

XX PR 15-SEP-1999; 99WO-US21090.

XX PR 15-SEP-1999; 99WO-US21547.

XX PR 08-OCT-1999; 99US-0158663.

XX PR 30-NOV-1999; 99WO-US28313.

XX PR 01-DEC-1999; 99WO-US28301.

XX PR 16-DEC-1999; 99WO-US30095.

XX PR 20-DEC-1999; 99WO-US30911.

XX PR 05-JAN-2000; 2000WO-US00219.

XX PR 06-JAN-2000; 2000WO-US00376.

XX PR 11-FEB-2000; 2000WO-US03565.

XX PR 18-FEB-2000; 2000WO-US04341.

XX PR 22-FEB-2000; 2000WO-US04414.

XX PR 24-FEB-2000; 2000WO-US04914.

XX PR 02-FEB-2000; 2000WO-US05004.

XX PR 02-MAR-2000; 2000WO-US05841.

XX PR 15-MAR-2000; 2000WO-US06884.

XX PR 20-MAR-2000; 2000WO-US07377.

(GETH ) GENENTECH INC.

XX Aahkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;

XX Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;

XX Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WL;

XX Zhang Z;

XX WPI; 2001-032160/04.

XX N-PSDB; AAF44115.

XX PRO polynucleotides used to produce polypeptides used to target

XX bioactive molecules such as toxins, radiolabels or antibodies, to

XX specific cells, to cause targeted cell death -

XX Claim 12; Fig 47; 935pp; English.

XX CC The present invention describes human secreted and transmembrane PRO  
proteins. The PRO proteins have cytostatic activity. The PRO proteins  
can be used for targeted delivery of bioactive molecules, such as  
toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
sequences, and their fragments, can be used as hybridisation probes, in  
chromosomal and gene mapping, and in the generation of anti-sense RNA





PA (MARS//) MARSTERS S A.  
PA (PANU//) PAN J.  
PA (PAON//) PAONI N F.  
PA (STEP//) STEPHAN J F.  
PA (WATA//) WATANABE C K.  
PA (WILL//) WILLIAMS P M.  
PA (WOOD//) WOOD W I.  
XX  
PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A,  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF,  
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
XX  
DR WPI; 2002-171999/22.  
DR N-PSDB; ABL95597.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal -  
XX  
PS Claim 11; Fig 74; 567pp; English.  
XX  
CC The present invention provides the protein and coding sequences of human  
CC PRO proteins. These are useful for treating or diagnosing a  
CC cardiovascular, endothelial or angiogenic disorder, including cardiac  
CC hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
XX healing. The present sequence is a PRO protein of the invention.  
XX  
SQ Sequence 867 AA;  
Query Match 99.4%; Score 4723.5; DB 23; Length 867;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 1 MGPPSLVLCCLASATVFSLLGSSAFLSHRLKGRFDRNRNIRNIIILVLTDDQVBLGS 60  
DB 1 MGPPSLVLCCLASATVFSLLGSSAFLSHRLKGRFDRNRNIRNIIILVLTDDQVBLGS 60  
QY 61 MQVANKTRIMEOGGAHFINAVTTPMCCPSRSSILTGKVVHNTYNNENCSPPSQA 120  
DB 61 MQVANKTRIMEOGGAHFINAVTTPMCCPSRSSILTGKVVHNTYNNENCSPPSQA 120  
QY 121 QHESRTFAVLNLTGYRTAFKGYLNEVNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180  
DB 121 QHESRTFAVLNLTGYRTAFKGYLNEVNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLITNDVSFRTSKOMYPHPPVLMVISHAAPHGPDSPQYSLRFP 240  
DB 181 EKHGSDYSKDYLTDLITNDVSFRTSKOMYPHPPVLMVISHAAPHGPDSPQYSLRFP 240  
QY 241 NASOHITPSYNAIPDKHIMRYTGMPKPIHMEFTNMLQKLOTLSVDDSMETIYNN 300  
DB 241 NASOHITPSYNAIPDKHIMRYTGMPKPIHMEFTNMLQKLOTLSVDDSMETIYNN 300  
QY 301 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360  
DB 301 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360  
QY 361 LNIDLAPTILDIAGLIDIPADMWGSILKLDTERPVNRFLKKQVRWDSFLVERGKLL 420  
DB 361 LNIDLAPTILDIAGLIDIPADMWGSILKLDTERPVNRFLKKQVRWDSFLVERGKLL 420  
QY 421 HKRNDKVDQEEENFLPKYQVVKDCQRAEYQTAECQLOKQWCEVDATCKLKHCKGP 480  
DB 421 HKRNDKVDQEEENFLPKYQVVKDCQRAEYQTAECQLOKQWCEVDATCKLKHCKGP 480  
QY 481 MRLGGSRAISNLVPKYTGQSGSEACTCSDGYKSLAGRRKLLFKKKYKASYVRSIRSV 540  
DB 481 MRLGGSRAISNLVPKYTGQSGSEACTCSDGYKSLAGRRKLLFKKKYKASYVRSIRSV 540  
QY 541 AIEVDGRVYHVGLGDAAPRNLTKRHWFGAPEDQDDKDGDFSTGGLPDYSAANPIKVT 600

DB 541 AIEVDGRVYHVGLGDAAPRNLTKRHWFGAPEDQDDKDGDFSTGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDLYKSLOAKWOKHLHDHETLQNKIKNREVGRHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDLDLYKSLOAKWOKHLHDHETLQNKIKNREVGRHLKKRPEEC 660  
QY 661 DCHKISYHTQHKRGLKRGSSSLHPPFKGLQEKDKVWLLREQKRRKKLKLRLQNNDT 720  
DB 661 DCHKISYHTQHKRGLKRGSSSLHPPFKGLQEKDKVWLLREQKRRKKLKLRLQNNDT 720  
QY 721 SMPEGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCMTINETHNPLFCFATGFL 780  
DB 721 SMPEGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCMTINETHNPLFCFATGFL 780  
QY 781 YFDLNTDPPQLMNAVNTLDRDVNLQHLVQLMELRSCKYKCNPRTRNDLGLKDGGSYE 840  
DB 781 YFDLNTDPPQLMNAVNTLDRDVNLQHLVQLMELRSCKYKCNPRTRNDLGLKDGGSYE 840  
QY 841 QYRQFQRRKWPENKRPSSKSLGQLMEGWEG 870  
DB 838 QYRQFQRRKWPENKRPSSKSLGQLMEGWEG 867  
RESULT 15  
ABB84853  
ID ABB84853 standard; Protein; 867 AA.  
XX  
AC ABB84853;  
XX  
DT 16-MAY-2002 (first entry)  
DE Human PRO1120 protein sequence SEQ ID NO:74.  
XX  
KW Human; angiogenesis; cardiant; cytostatic; angiogenic; hypotensive;  
KW vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;  
KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;  
KW angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;  
KW age-related macular degeneration; arterial restenosis; angina;  
KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;  
KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;  
KW wound healing; chromosome mapping; gene mapping.  
OS Homo sapiens.  
PN WO200200690-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 20-JUN-2001; 2001WO-US19692.  
XX  
PR 23-JUN-2000; 2000US-213637P.  
PR 20-JUL-2000; 2000US-219556P.  
PR 25-JUL-2000; 2000US-220624P.  
PR 25-JUL-2000; 2000US-220664P.  
PR 28-JUL-2000; 2000US-2020710.  
PR 02-AUG-2000; 2000US-222695P.  
PR 17-AUG-2000; 2000US-0643657.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 07-SEP-2000; 2000US-230978P.  
PR 18-SEP-2000; 2000US-064610.  
PR 18-SEP-2000; 2000US-0665350.  
PR 24-OCT-2000; 2000US-242922P.  
PR 08-NOV-2000; 2000US-0709238.  
PR 08-NOV-2000; 2000WO-US10952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000US-0747259.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 22-JAN-2001; 2001US-0767609.  
PR 28-FEB-2001; 2001US-0796498.  
PR 28-FEB-2001; 2001WO-US06520.

PR 01-MAR-2001; 2001WO-US06666.  
 PR 09-MAR-2001; 2001US-0802706.  
 PR 14-MAR-2001; 2001US-0808689.  
 PR 22-MAR-2001; 2001US-0816744.  
 PR 05-APR-2001; 2001US-0828366.  
 PR 10-MAY-2001; 2001US-0854208.  
 PR 10-MAY-2001; 2001US-0854280.  
 PR 25-MAY-2001; 2001US-0866028.  
 PR 25-MAY-2001; 2001US-0866034.  
 PR 25-MAY-2001; 2001WO-US17092.  
 PR 30-MAY-2001; 2001US-0870574.  
 PR 30-MAY-2001; 2001WO-US17443.  
 PR 01-JUN-2001; 2001WO-US17800.  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
 PI Godowski PJ, Gurney AL, Hillan KJ, Masters SA, Pan J, Paoni NF;  
 PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
 XX  
 XX WPI: 2002-090516/12.  
 DR N-PSDB; ABJ88108.  
 DR  
 XX  
 XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
 PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
 PT infarction), endothelial or angiogenic disorders in a mammal -  
 XX  
 XX Claim 11; Fig 74; 565pp; English.  
 XX  
 CC ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to  
 CC ABB85003. The PRO proteins and polynucleotides have cardiac, cytostatic,  
 CC antiangiogenic, hypotensive, vulnary and antiarteriosclerotic  
 CC activities, and can be used in gene therapy. The PRO polynucleotides,  
 CC proteins, agonists and antagonists are useful for treating or diagnosing  
 CC a cardiovascular, endothelial or angiogenic disorder in a mammal,  
 CC e.g. cardiac hypertrophy, trauma, cancer, age-related macular  
 CC degeneration, atherosclerosis, hypertension, arterial restenosis,  
 CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis,  
 CC lymphangitis, tumour angiogenesis (such as breast carcinoma and liver  
 CC carcinoma) and wound healing. The PRO polynucleotides have applications  
 CC in molecular biology, including use as hybridisation probes, and in  
 CC chromosome and gene mapping. ABL88259 to ABL88267 represent primers and  
 CC probes used in the exemplification of the present invention.  
 XX  
 XX Sequence 867 AA;  
 SQ  
 Query Match 99.4%; Score 4723.5; DB 23; Length 867;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
 QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
 Db 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
 QY 61 MOVNKRTRIMEGGGAHFNAVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120  
 Db 61 MOVNKRTRIMEGGGAHFNAVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120  
 QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180  
 Db 121 QHESRTFAVYLNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180  
 QY 181 EKHGSDYSKDYLTDLITNDSVFFRTSKMTPHRPVLVMI SHAAPHGPEDSAPQYSRLFP 240  
 Db 181 EKHGSDYSKDYLTDLITNDSVFFRTSKMTPHRPVLVMI SHAAPHGPEDSAPQYSRLFP 240  
 QY 241 NASQHITPSYNAFPDKHWIMRYTGPMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNN 300  
 Db 241 NASQHITPSYNAFPDKHWIMRYTGPMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNN 300  
 QY 301 LVETGELDNTIYVTADHGHIHQGLVKGSMPYEFDIRVPPYVGRPNVEAGCLNPHIV 360  
 Db 301 LVETGELDNTIYVTADHGHIHQGLVKGSMPYEFDIRVPPYVGRPNVEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRPHLKKKMRVWRDSFLVERGKLL 420  
 Db 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRPHLKKKMRVWRDSFLVERGKLL 420  
 QY 421 HKDNDKVDQAQENFLPKYQVRVKDLQORABYOTACEQLQKQKQCVEDATGKLHKCKGP 480  
 Db 421 HKDNDKVDQAQENFLPKYQVRVKDLQORABYOTACEQLQKQKQCVEDATGKLHKCKGP 480  
 QY 481 MRLGGSRLSNLVPKYGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYVRSRISRV 540  
 Db 481 MRLGGSRLSNLVPKYGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYVRSRISRV 540  
 QY 541 AIEVDGRVHVGLGDAQAQPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Db 541 AIEVDGRVHVGLGDAQAQPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 QY 601 HRCYIENDTVQCDLQKSLQAKDKHLHDIETLQNKILREVRGHLKKRPEEC 660  
 Db 601 HRCYIENDTVQCDLQKSLQAKDKHLHDIETLQNKILREVRGHLKKRPEEC 660  
 QY 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYWCMTINETHNFI FCEFATGPLE 780  
 Db 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYWCMTINETHNFI FCEFATGPLE 780  
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 QY 841 QYRQFORRKWPEWKRPSKSLGOLWEGWEG 870  
 Db 838 QYRQFORRKWPEWKRPSKSLGOLWEGWEG 867

Search completed: February 14, 2004, 22:54:05  
 Job time : 76 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 19:29:52 ; Search time 1133 Seconds  
(without alignments)  
10933.854 Million cell updates/sec

Title: US-10-025-966A-4  
Perfect score: 3363  
Sequence: 1 gactccgcgtcccaaga.....aaagaaaaaa 3363

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3363	100.0	3363	13	US-10-025-966A-4
2	3363	100.0	3363	13	US-10-265-071-4
3	3348.4	99.6	4303	12	US-10-264-237-1349
4	3347.2	99.5	4397	13	US-09-970-287-11
5	3346.8	99.5	4286	13	US-10-025-966A-13
6	3346.8	99.5	4286	13	US-10-265-071-13
7	3344.8	99.5	3348	13	US-10-182-951-17
8	3323.4	98.8	3906	9	US-09-989-722-83
9	3323.4	98.8	3906	9	US-09-989-723-83
10	3323.4	98.8	3906	9	US-09-989-727-83
11	3323.4	98.8	3906	9	US-09-989-729-83
12	3323.4	98.8	3906	10	US-09-989-731-83
13	3323.4	98.8	3906	10	US-09-989-732-83
14	3323.4	98.8	3906	10	US-09-991-073-83
15	3323.4	98.8	3906	10	US-09-990-442-83

16	3323.4	98.8	3906	10	US-09-991-163-83	Sequence 83, Appl
17	3323.4	98.8	3906	10	US-09-993-604-83	Sequence 83, Appl
18	3323.4	98.8	3906	10	US-09-990-456-83	Sequence 83, Appl
19	3323.4	98.8	3906	10	US-09-989-721-83	Sequence 83, Appl
20	3323.4	98.8	3906	10	US-09-992-598-83	Sequence 83, Appl
21	3323.4	98.8	3906	10	US-09-989-293A-83	Sequence 83, Appl
22	3323.4	98.8	3906	10	US-09-989-735-83	Sequence 83, Appl
23	3323.4	98.8	3906	10	US-09-990-444-83	Sequence 83, Appl
24	3323.4	98.8	3906	10	US-09-991-181-83	Sequence 83, Appl
25	3323.4	98.8	3906	10	US-09-989-730-83	Sequence 83, Appl
26	3323.4	98.8	3906	10	US-09-990-436-83	Sequence 83, Appl
27	3323.4	98.8	3906	10	US-09-993-687-83	Sequence 83, Appl
28	3323.4	98.8	3906	11	US-09-989-734-83	Sequence 83, Appl
29	3323.4	98.8	3906	11	US-09-997-653-83	Sequence 83, Appl
30	3323.4	98.8	3906	11	US-09-993-667-83	Sequence 83, Appl
31	3323.4	98.8	3906	11	US-09-997-428-83	Sequence 83, Appl
32	3323.4	98.8	3906	11	US-09-997-666-83	Sequence 83, Appl
33	3323.4	98.8	3906	11	US-09-990-438-83	Sequence 83, Appl
34	3323.4	98.8	3906	11	US-09-990-562-83	Sequence 83, Appl
35	3323.4	98.8	3906	11	US-09-990-711-83	Sequence 83, Appl
36	3323.4	98.8	3906	11	US-09-989-726-83	Sequence 83, Appl
37	3323.4	98.8	3906	11	US-09-998-156-83	Sequence 83, Appl
38	3323.4	98.8	3906	11	US-09-990-437-83	Sequence 83, Appl
39	3323.4	98.8	3906	11	US-09-991-157-83	Sequence 83, Appl
40	3323.4	98.8	3906	11	US-09-997-514-83	Sequence 83, Appl
41	3323.4	98.8	3906	11	US-09-997-573-83	Sequence 83, Appl
42	3323.4	98.8	3906	11	US-09-991-172-83	Sequence 83, Appl
43	3323.4	98.8	3906	11	US-09-990-726-83	Sequence 83, Appl
44	3323.4	98.8	3906	11	US-09-997-559-83	Sequence 83, Appl
45	3323.4	98.8	3906	11	US-09-997-601-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1  
US-10-025-966A-4  
; Sequence 4, Application US/10025966A  
; Publication No. US20030148920A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Steven D.  
; APPLICANT: Palmeri, Diana  
; APPLICANT: Stefan, Hemmerich  
; TITLE OF INVENTION: Sulfatases and methods of use  
; FILE REFERENCE: UCAL230  
; CURRENT APPLICATION NUMBER: US/10/025.966A  
; CURRENT FILING DATE: 2002-12-21  
; PRIOR APPLICATION NUMBER: 60/258,577  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 60/267,831  
; PRIOR FILING DATE: 2001-09-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 3363  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-025-966A-4

Query Match	100.0%;	Score 3363;	DB 13;	Length 3363;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3363;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GACTCCGCATCCCAAGAACGACAGATCAGCAAAAAGAGATGGGCCCCCGAGC	60	
Db	1	GACTCCGCATCCCAAGAACGACAGATCAGCAAAAAGAGATGGGCCCCCGAGC	60	
QY	61	CTCGTGCTGTGCTTCTGTCGCAACTGTGTTCCTCTGCTGGTGAAGCTCGGCTTC	120	
Db	61	CTCGTGCTGTGCTTCTGTCGCAACTGTGTTCCTCTGCTGGTGAAGCTCGGCTTC	120	
QY	121	CTGTGCGACACCCCTGAAAGGAGGAGGTTTCAGAGGAGCGGACATCCGCCCCAAC	180	

121 CTGTGCGCACCGCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCAAC 180  
181 ATCATCTCTGTGCTGACGCGACGACGAGATGTGAGCTGGGTTCCATGCGAGGTATGAAC 240  
181 ATCATCTCTGTGCTGACGCGACGACGAGATGTGAGCTGGGTTCCATGCGAGGTATGAAC 240  
241 AAGACCCCGCGCATCATGGAGCAGGCGGGGGCGCACTTCATCAACGCGCTTCGTGACCACA 300  
241 AAGACCCCGCGCATCATGGAGCAGGCGGGGGCGCACTTCATCAACGCGCTTCGTGACCACA 300  
301 CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGCGAGTACGTCCCAACCAAC 360  
301 CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGCGAGTACGTCCCAACCAAC 360  
361 ACCTACACCAACATGAGAACTGTCTCTCGCCCTCTGCGAGCAGCAGCAGAGCGCG 420  
361 ACCTACACCAACATGAGAACTGTCTCTCGCCCTCTGCGAGCAGCAGCAGAGCGCG 420  
421 ACCTTTCGGTGTACTCTATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480  
421 ACCTTTCGGTGTACTCTATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480  
481 AATGAATACACGCGCTCTACGTGCGACCGCGCTGGAAGAGTGGTCCGACTCCTTAAA 540  
481 AATGAATACACGCGCTCTACGTGCGACCGCGCTGGAAGAGTGGTCCGACTCCTTAAA 540  
541 AACTCCGCTTTTAACTAATACGCTGTGTGCGAAACGGGGTGAAGAAAGCAGCGCTCC 600  
541 AACTCCGCTTTTAACTAATACGCTGTGTGCGAAACGGGGTGAAGAAAGCAGCGCTCC 600  
601 GACTACTCCAAAGATTAATACAGACTCTATCAACCAATGACAGCTGAGCTTCTCCGC 660  
601 GACTACTCCAAAGATTAATACAGACTCTATCAACCAATGACAGCTGAGCTTCTCCGC 660  
661 ACCTCCAAAGATTAATACAGACTCTATCAACCAATGACAGCTGAGCTTCTCCGC 720  
661 ACCTCCAAAGATTAATACAGACTCTATCAACCAATGACAGCTGAGCTTCTCCGC 720  
721 CAGGCGCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAGCGATCTCAGGAC 780  
721 CAGGCGCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAGCGATCTCAGGAC 780  
781 ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAACTGGATCATCGCTACACG 840  
781 ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAACTGGATCATCGCTACACG 840  
841 GGGCCCATGAAGCCCATCCATGAAATTCACCAATGCTCCAGCGGAAGCGCTTGCGAG 900  
841 GGGCCCATGAAGCCCATCCATGAAATTCACCAATGCTCCAGCGGAAGCGCTTGCGAG 900  
901 ACCCTCATGTCGGTGGACGACTCCATGGAGAGATTTACAACTGCTGGTTGAGACGGGC 960  
901 ACCCTCATGTCGGTGGACGACTCCATGGAGAGATTTACAACTGCTGGTTGAGACGGGC 960  
961 GAGCTGGACACACGTPACATGCTATACACCGCCGACCAAGTTACCACTCGGCAGTTT 1020  
961 GAGCTGGACACACGTPACATGCTATACACCGCCGACCAAGTTACCACTCGGCAGTTT 1020  
1021 GGCCTGGTGAAGGGAAATCCATGCGATATGAGTTTGAATCATCAGGCTCCGTTCTACGTG 1080  
1021 GGCCTGGTGAAGGGAAATCCATGCGATATGAGTTTGAATCATCAGGCTCCGTTCTACGTG 1080  
1081 AGGGGCCCCAAGTGGAAAGCGGCTGTCTGAATCCCCCACTGCTCTCAACATTTGACCTG 1140  
1081 AGGGGCCCCAAGTGGAAAGCGGCTGTCTGAATCCCCCACTGCTCTCAACATTTGACCTG 1140  
1141 GCGCCACCATCTGTGACATTTGAGGCTTGGACATACCTGCGGATATGAGCGGAATTC 1200  
1141 GCGCCACCATCTGTGACATTTGAGGCTTGGACATACCTGCGGATATGAGCGGAATTC 1200  
1201 ATCTCTCAAGCTGTGACACGAGCGCGCGGTGAATCGTTTCACTTCAAAAAAGAGATG 1260

1201 ATCTCTCAAGCTGTGACACGAGCGCGCGGTGAATCGGTTTCACTTGA AAAAAGAGATG 1260  
1261 AGGGTCTGCGGGACTCTCTTTTGTGTGAGAGAGGCAAGCTCTACACAAGAGAGACAAT 1320  
1261 AGGGTCTGCGGGACTCTCTTTTGTGTGAGAGAGGCAAGCTCTACACAAGAGAGACAAT 1320  
1321 GACAAGGTGGACGCCGACGAGGAGAACTTTTCTGCCCAAGTACCAAGCTGTGAAGACCTG 1380  
1321 GACAAGGTGGACGCCGACGAGGAGAACTTTTCTGCCCAAGTACCAAGCTGTGAAGACCTG 1380  
1381 TGTCAAGCTGTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAAGTGGCAAGTGTG 1440  
1381 TGTCAAGCTGTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAAGTGGCAAGTGTG 1440  
1441 GAGGACGCCACGCGGAGAGCTGAAGCTGCAATAGTCAAGGGGCCCATGCGGTGGGCGGC 1500  
1441 GAGGACGCCACGCGGAGAGCTGAAGCTGCAATAGTCAAGGGGCCCATGCGGTGGGCGGC 1500  
1501 AGCAGAGCCCTCTCCAACTCTGCTGCCCAAGTACTACGGGACGGGACGAGGCTGCAAC 1560  
1501 AGCAGAGCCCTCTCCAACTCTGCTGCCCAAGTACTACGGGACGGGACGAGGCTGCAAC 1560  
1561 TGTGACAGCGGGGACTTACAAAGCTCAGCCCTGGCGGACGCCGGA AAAA ACTCTTCAAGAG 1620  
1561 TGTGACAGCGGGGACTTACAAAGCTCAGCCCTGGCGGACGCCGGA AAAA ACTCTTCAAGAG 1620  
1621 AAGTACAGGCCAGCTATGTCCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 1680  
1621 AAGTACAGGCCAGCTATGTCCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 1680  
1681 GGCAGGGTGTACCAAGTGGCTGATGCGCGCCAGCGCCGGA AAAA CCTCAAGAGCGG 1740  
1681 GGCAGGGTGTACCAAGTGGCTGATGCGCGCCAGCGCCGGA AAAA CCTCAAGAGCGG 1740  
1741 CACTGCGCAGGGGCCCCGTGAGCAAGATGACAAAGATGAGTGGGGA CTTCAGTGGCACT 1800  
1741 CACTGCGCAGGGGCCCCGTGAGCAAGATGACAAAGATGAGTGGGGA CTTCAGTGGCACT 1800  
1801 GAGAGCTTCCGAGTACTCTCAGCGCCCAACCCATTAAAGTGAACATCGGTGCTACATC 1860  
1801 GAGAGCTTCCGAGTACTCTCAGCGCCCAACCCATTAAAGTGAACATCGGTGCTACATC 1860  
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGCAGGCTGGAAA 1920  
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGCAGGCTGGAAA 1920  
1921 GACCAACAGCTGCAATGCAACCAAGATGAAACCTTGCAGAACAAAATTAAGAACCTG 1980  
1921 GACCAACAGCTGCAATGCAACCAAGATGAAACCTTGCAGAACAAAATTAAGAACCTG 1980  
1981 AGGGAAGTCCGAGGTCACCTGAAAGAAAAGCGGCCAGAGAAATGTGACTGCACAAAATC 2040  
1981 AGGGAAGTCCGAGGTCACCTGAAAGAAAAGCGGCCAGAGAAATGTGACTGCACAAAATC 2040  
2041 AGCTACACACCCAGGACAAAAGGCGCTCAAGACACAGGCTCCAGTCTGCATCTCTTTC 2100  
2041 AGCTACACACCCAGGACAAAAGGCGCTCAAGACACAGGCTCCAGTCTGCATCTCTTTC 2100  
2101 AGGAAGGGCTGCAAGAGAAAGGACAGGTGTCGCTTTTSCGGGAGCAGAGCGCAAGAG 2160  
2101 AGGAAGGGCTGCAAGAGAAAGGACAGGTGTCGCTTTTSCGGGAGCAGAGCGCAAGAG 2160  
2161 AAACTCCGCAAGCTGCTCAAGCGCTTGCAGAACAAACGACACGTCAGCATGCCGCTTC 2220  
2161 AAACTCCGCAAGCTGCTCAAGCGCTTGCAGAACAAACGACACGTCAGCATGCCGCTTC 2220  
2221 AGTGTCTTCAACCAAGCACTGGGACAGCGGCTTTTCTGGACACTGGGGCT 2280  
2221 AGTGTCTTCAACCAAGCACTGGGACAGCGGCTTTTCTGGACACTGGGGCT 2280  
2281 TTCTGTGCTGACACGAGCCCAACCAATTAACAGTACTGGTGCATGAGGACCAATCAATGAG 2340  
2281 TTCTGTGCTGACACGAGCCCAACCAATTAACAGTACTGGTGCATGAGGACCAATCAATGAG 2340





Db 601 GACTACTCCAGGATTACCTCAGAGACTCATCAACCAATGACAGCGTGAAGCTTTCCCGC 660  
Qy 661 ACCTCCAGAGATGTATCCGCCACAGGCCAGTCTCTCATGTGTATCATGACCAATGACGCCCTCC 720  
Db 661 ACCTCCAGAGATGTATCCGCCACAGGCCAGTCTCTCATGTGTATCATGACCAATGACGCCCTCC 720  
Qy 721 CACGCCCTGAGGATTCAGCCCAATATATCAAGCTCTCTCCAAACGATCTCAGCAC 780  
Db 721 CACGCCCTGAGGATTCAGCCCAATATATCAAGCTCTCTCCAAACGATCTCAGCAC 780  
Qy 781 ATACGCCGAGCTACAACTAGCGCCCAACCCCGGACAAACATCTGGATCATGCGCTACACG 840  
Db 781 ATACGCCGAGCTACAACTAGCGCCCAACCCCGGACAAACATCTGGATCATGCGCTACACG 840  
Qy 841 GGGCCCATGAGCCATTCACATGGAATTCACCAACATGCTCAGCGGAGCGCTTGCAG 900  
Db 841 GGGCCCATGAGCCATTCACATGGAATTCACCAACATGCTCAGCGGAGCGCTTGCAG 900  
Qy 901 ACCCTCATGTGCTGGAGAGCTCCATCGAGACGATTTACAAATGCTGGTTGAGACGGGC 960  
Db 901 ACCCTCATGTGCTGGAGAGCTCCATCGAGACGATTTACAAATGCTGGTTGAGACGGGC 960  
Qy 961 GAGCTGACACACGATCATGCTATACACCGCCGACCAACGTTACCAATCGGCCAGTTT 1020  
Db 961 GAGCTGACACACGATCATGCTATACACCGCCGACCAACGTTACCAATCGGCCAGTTT 1020  
Qy 1021 GGCCTGTGAAAGGGAATCCATGCCATATAGTTTGAATCATGAGGTCCGCTTACAGTG 1080  
Db 1021 GGCCTGTGAAAGGGAATCCATGCCATATAGTTTGAATCATGAGGTCCGCTTACAGTG 1080  
Qy 1081 AGGGCCCCAAGCTGGAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1140  
Db 1081 AGGGCCCCAAGCTGGAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTGACCTG 1140  
Qy 1141 GCCCCACCATCTCGGACATTCAGAGCTGGACATACCTCGGATATGACCGGAAATCC 1200  
Db 1141 GCCCCACCATCTCGGACATTCAGAGCTGGACATACCTCGGATATGACCGGAAATCC 1200  
Qy 1201 ATCTCAAGCTGTGGACACGAGCGGCTGTGAATCGGTTTCACTTGAAGAAAGATG 1260  
Db 1201 ATCTCAAGCTGTGGACACGAGCGGCTGTGAATCGGTTTCACTTGAAGAAAGATG 1260  
Qy 1261 AGGCTCTGGCGGACCTCTTTGGTGAGAGAGGCAAGCTGCTCAACAGAGACAT 1320  
Db 1261 AGGCTCTGGCGGACCTCTTTGGTGAGAGAGGCAAGCTGCTCAACAGAGACAT 1320  
Qy 1321 GACAAGTGGACGCCAGGAGGAGAACTTTCTGCCAAGTACAGAGCTGTGAAGACCTG 1380  
Db 1321 GACAAGTGGACGCCAGGAGGAGAACTTTCTGCCAAGTACAGAGCTGTGAAGACCTG 1380  
Qy 1381 TGTACGCTGTGATGATACAGAGCGGCTGTGAGAGCTGGGACAGAAAGTGGCAGTGTG 1440  
Db 1381 TGTACGCTGTGATGATACAGAGCGGCTGTGAGAGCTGGGACAGAAAGTGGCAGTGTG 1440  
Qy 1441 GAGGACGCCACGGGAGCTGAAGCTGCAATAGTGAAGGCCCAATCGGCTGGCGCGC 1500  
Db 1441 GAGGACGCCACGGGAGCTGAAGCTGCAATAGTGAAGGCCCAATCGGCTGGCGCGC 1500  
Qy 1501 AGCAGAGCTCTTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGAGCCCTGCACC 1560  
Db 1501 AGCAGAGCTCTTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGAGCCCTGCACC 1560  
Qy 1561 TGTGACAGCGGAGCTACAGCTCAGCTGCGCGGACCGCGGAAACCTCTCAAGAG 1620  
Db 1561 TGTGACAGCGGAGCTACAGCTCAGCTGCGCGGACCGCGGAAACCTCTCAAGAG 1620  
Qy 1621 AAGTACAAGCCAGCTATGCTCGCAGTCTCATCGCTCAGTGGCCATCAGGTGGAC 1680  
Db 1621 AAGTACAAGCCAGCTATGCTCGCAGTCTCATCGCTCAGTGGCCATCAGGTGGAC 1680  
Qy 1681 GGCAGGCTGACACGATGAGCTTGGTGTATGCGGCCCGGAAACCTCAACAGCGG 1740  
Db 1681 GGCAGGCTGACACGATGAGCTTGGTGTATGCGGCCCGGAAACCTCAACAGCGG 1740

Qy 1741 CACTGGCCAGGGCCCTTGAGGACCAAGATGACAGGATGTTGGGAGCTTCAGTGGCACT 1800  
Db 1741 CACTGGCCAGGGCCCTTGAGGACCAAGATGACAGGATGTTGGGAGCTTCAGTGGCACT 1800  
Qy 1801 GGAGGCTTCCGACTACTAGCGCCCAACCCCAATTAAGTGAACATCGTGTCTACATC 1860  
Db 1801 GGAGGCTTCCGACTACTAGCGCCCAACCCCAATTAAGTGAACATCGTGTCTACATC 1860  
Qy 1861 CTAGAGAACGACACAGTCCAGTCCAGTGTGACCTGTACAGTCCCTGCGAGCCCTGAAA 1920  
Db 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGTGACCTGTACAGTCCCTGCGAGCCCTGAAA 1920  
Qy 1921 GACCAACAGCTGCACATGACCCAGAGATTGAACCCCTGCAGAACAAAAATTAAGAACCTG 1980  
Db 1921 GACCAACAGCTGCACATGACCCAGAGATTGAACCCCTGCAGAACAAAAATTAAGAACCTG 1980  
Qy 1981 AGGGAAGTCCGAGTCCAGTCCAGTGAAGAAAGGCGCCAGAGAAATGTGACTGTCAAAAATC 2040  
Db 1981 AGGGAAGTCCGAGTCCAGTGAAGAAAGGCGCCAGAGAAATGTGACTGTCAAAAATC 2040  
Qy 2041 AGCTACCAACCCAGACAAAGGCGCTCAAGCAGAGAGCTCCAGTCTGCACTCTTTC 2100  
Db 2041 AGCTACCAACCCAGACAAAGGCGCTCAAGCAGAGAGCTCCAGTCTGCACTCTTTC 2100  
Qy 2101 AGGAGGCGCTGCAAGAGAAAGGCAAGGTGTGGTGTGGGAGCGAAGCGCAAGAG 2160  
Db 2101 AGGAGGCGCTGCAAGAGAAAGGCAAGGTGTGGTGTGGGAGCGAAGCGCAAGAG 2160  
Qy 2161 AAACCTCGCAAGCTGTCAAGCGCTCGAAGAAACAAACGACACGTGCGAGCTCCAGAGGCTC 2220  
Db 2161 AAACCTCGCAAGCTGTCAAGCGCTCGAAGAAACAAACGACACGTGCGAGCTCCAGAGGCTC 2220  
Qy 2221 ACCTGCTTCAACCAAGCAACCGAGCTGCGAGAGCGGCTTTCTGGACACTGGGCGCT 2280  
Db 2221 ACCTGCTTCAACCAAGCAACCGAGCTGCGAGAGCGGCTTTCTGGACACTGGGCGCT 2280  
Qy 2281 TTCTGTGCTGCAACCGCGCAACAAATACACGTACTGTGTCATGAGGACCATCAATGAG 2340  
Db 2281 TTCTGTGCTGCAACCGCGCGCAACAAATACACGTACTGTGTCATGAGGACCATCAATGAG 2340  
Qy 2341 ACTCAAAATTTCTTCTGTGAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 2400  
Db 2341 ACTCAAAATTTCTTCTGTGAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 2400  
Qy 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGAGAGAGGATGCTCTCAACAG 2460  
Db 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGAGAGAGGATGCTCTCAACAG 2460  
Qy 2461 CTACACGTCATGAGCTGAGAGCTGCAAGGCTTACAGAGCTGTAACCCCGG 2520  
Db 2461 CTACACGTCATGAGCTGAGAGCTGCAAGGCTTACAGAGCTGTAACCCCGG 2520  
Qy 2521 ACTGGAACATGACCTGGGACTTAAAGATGAGAGGAGCTATGAGCAATACAGGCGATTT 2580  
Db 2521 ACTGGAACATGACCTGGGACTTAAAGATGAGAGGAGCTATGAGCAATACAGGCGATTT 2580  
Qy 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Db 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Qy 2641 GAAGGCTGGGAAGGTTAAGAAACAAAGAGGTTGGAACCTCCAAAAACATAGAGCATCAAC 2700  
Db 2641 GAAGGCTGGGAAGGTTAAGAAACAAAGAGGTTGGAACCTCCAAAAACATAGAGCATCAAC 2700  
Qy 2701 TGACTGCACAGGCAATGAAGAAACCATGTGGTGTATTTCCAGGACAGCTGTGCTATTGGCC 2760  
Db 2701 TGACTGCACAGGCAATGAAGAAACCATGTGGTGTATTTCCAGGACAGCTGTGCTATTGGCC 2760  
Qy 2761 AGGAGGCTTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCTGGAGGATAAC 2820  
Db 2761 AGGAGGCTTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCTGGAGGATAAC 2820

QY 2821 CAGCAGGAGCAGATACCTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
DB |||||  
QY 2821 CAGCAGGAGCAGATACCTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
DB |||||  
QY 2881 TACCTCACCAGCTGCACAAAATGCAATTTTCGATCAAAAAGTCCACCACTAACCTCC 2940  
DB |||||  
QY 2881 TACCTCACCAGCTGCACAAAATGCAATTTTCGATCAAAAAGTCCACCACTAACCTCC 2940  
DB |||||  
QY 2941 CCAGAAGCTCAAAAGGAAACCGAGAGAGCGAGAGAGATTTCTTTGGAAATTTCT 3000  
DB |||||  
QY 3001 CCCAAGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCACTCTCTTCTAAA 3060  
DB |||||  
QY 3001 CCCAAGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCACTCTCTTCTAAA 3060  
DB |||||  
QY 3061 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGAACTTAAGAAGCAGACAGAGGCAACGT 3120  
DB |||||  
QY 3061 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGAACTTAAGAAGCAGACAGAGGCAACGT 3120  
DB |||||  
QY 3121 GGAGAGCTGAAGACAGTGCAGAGACGTTTGAATAGTCAATAGTCAAGAGAGATGA 3180  
DB |||||  
QY 3121 GGAGAGCTGAAGACAGTGCAGAGACGTTTGAATAGTCAATAGTCAAGAGAGATGA 3180  
DB |||||  
QY 3181 CATTTACCTAGCACTATAAAACCCCTGGTTGCTCTGAAGAACTGCCTTCAATGATATAT 3240  
DB |||||  
QY 3181 CATTTACCTAGCACTATAAAACCCCTGGTTGCTCTGAAGAACTGCCTTCAATGATATAT 3240  
DB |||||  
QY 3241 GTGACTATTACATGTAATCAACATGGAACTTTTAGGGAACTTAAGAATAATCCCAA 3300  
DB |||||  
QY 3241 GTGACTATTACATGTAATCAACATGGAACTTTTAGGGAACTTAAGAATAATCCCAA 3300  
DB |||||  
QY 3301 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGAAGAAAGAAABAAA 3360  
DB |||||  
QY 3301 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGAAGAAAGAAABAAA 3360  
DB |||||  
QY 3361 AAA 3363  
DB |||||  
QY 3361 AAA 3363  
DB |||||

## RESULT 3

US-10-264-237-1349  
; Sequence 1349, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birex et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P413P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 1349  
; LENGTH: 4303  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (24)..(24)  
; OTHER INFORMATION: n equals a,t,g, or c

Query Match

Best Local Similarity

99.6%; Score 3348.4; DB 12; Length 4303;  
99.9%; Pred. No. 0;

Matches 3349; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GACTCCCGCATCCCAAAAGAGCACCACATCAGCAAAAAGAGAGATGGCCCCCGAGC 60  
DB |||||  
QY 177 GACTCCCGCATCCCAAAAGAGCACCACATCAGCAAAAAGAGAGATGGCCCCCGAGC 236  
DB |||||  
QY 61 CTCGTGCTGTGCTGTCTCCGCAACTGTCTCTCCCTGCTGGGTGGAGCTCGSCCTTC 120  
DB |||||  
QY 237 CTCGTGCTGTGCTGTCTCCGCAACTGTCTCTCCCTGCTGGGTGGAGCTCGSCCTTC 296  
DB |||||  
QY 121 CTGTGCGCACCCCGCTCAAAAGGAGGTTTCAGAGGGACCGCAGGAAACATCGCCCCAAC 180  
DB |||||  
QY 297 CTGTGCGCACCCCGCTCAAAAGGAGGTTTCAGAGGGACCGCAGGAAACATCGCCCCAAC 356  
DB |||||  
QY 181 ATCATCTCTGGTGTGACGAGACGACAGGATGTGGAGCTGGGTTCATCAGGTGATGAAC 240  
DB |||||  
QY 357 ATCATCTCTGGTGTGACGAGACGACAGGATGTGGAGCTGGGTTCATCAGGTGATGAAC 416  
DB |||||  
QY 241 AAGACCCGCGCATCATGAGAGGCGGGGCGCACTTCATCAACGGCTTCGTGACACA 300  
DB |||||  
QY 417 AAGACCCGCGCATCATGAGAGGCGGGGCGCACTTCATCAACGGCTTCGTGACACA 476  
DB |||||  
QY 301 CCCATGTCTGCCCTCCTCAGCTCTCCATCTCTGCAAGTACGTCCACAAACCAAC 360  
DB |||||  
QY 477 CCCATGTCTGCCCTCCTCAGCTCTCCATCTCTCAGGCAAGTACGTCCACAAACCAAC 536  
DB |||||  
QY 361 ACCTACACCAACATAGAACTGTCTCTCGCCCTCTGGCAGGACAGCAGAGCGCG 420  
DB |||||  
QY 537 ACCTACACCAACATAGAACTGTCTCTCGCCCTCTGGCAGGACAGCAGAGCGCG 596  
DB |||||  
QY 421 ACCTTTGCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 480  
DB |||||  
QY 597 ACCTTTGCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 656  
DB |||||  
QY 481 AATGAATACAAACGGCTCTACGTGCCAACCCGGCTGGAAAGAGTGGGTGGACTCCTTAAA 540  
DB |||||  
QY 657 AATGAATACAAACGGCTCTACGTGCCAACCCGGCTGGAAAGAGTGGGTGGACTCCTTAAA 716  
DB |||||  
QY 541 AACTCCCGCTTTTATTAATACTACAGCTGTGTGCGAACGGGGTGAAGAAAGAGCGCTCC 600  
DB |||||  
QY 717 AACTCCCGCTTTTATTAATACTACAGCTGTGTGCGAACGGGGTGAAGAAAGAGCGCTCC 776  
DB |||||  
QY 601 GACTACTCCAAAGGATTACCTCAGAGACTCATCAACCAATGACAGCTGAGCTTCTTCGCG 660  
DB |||||  
QY 777 GACTACTCCAAAGGATTACCTCAGAGACTCATCAACCAATGACAGCTGAGCTTCTTCGCG 836  
DB |||||  
QY 661 ACGTCCAAAGAGATGTACCCGACAGGCCAGTCTCATGTGTCATCAGCCATGAGCCCC 720  
DB |||||  
QY 837 ACGTCCAAAGAGATGTACCCGACAGGCCAGTCTCATGTGTCATCAGCCATGAGCCCC 896  
DB |||||  
QY 721 CAGGCGCTGAGGATTACGCCCCCAATATTACGCTCTTCCCAACGCAATCTCAGCAC 780  
DB |||||  
QY 897 CAGGCGCTGAGGATTACGCCCCCAATATTACGCTCTTCCCAACGCAATCTCAGCAC 956  
DB |||||  
QY 781 ATCAGCGCGAGCTACAACTAGCGCCCAACCCGACAAACACATGGATCATGCGTCAACG 840  
DB |||||  
QY 957 ATCAGCGCGAGCTACAACTAGCGCCCAACCCGACAAACACATGGATCATGCGTCAACG 1016  
DB |||||  
QY 841 GGGCCCATGAGCCCATCCCATGGAATTCACCAACATGCTCCAGCGGAGAGCGCTTCAG 900  
DB |||||  
QY 1017 GGGCCCATGAGCCCATCCCATGGAATTCACCAACATGCTCCAGCGGAGAGCGCTTCAG 1076  
DB |||||  
QY 901 ACCCTCATGTGCGTGGACGACTCCATGAGACGATTTTCAACATGCTGTTGGAGCGGGC 960  
DB |||||  
QY 1077 ACCCTCATGTGCGTGGACGACTCCATGAGACGATTTTCAACATGCTGTTGGAGCGGGC 1136  
DB |||||  
QY 961 GAGCTGGACACAGTATCATGCTATACACCGCCGACCAACGTTTACCAATCGGCCAGTTT 1020  
DB |||||  
QY 1137 GAGCTGGACACAGTATCATGCTATACACCGCCGACCAACGTTTACCAATCGGCCAGTTT 1196  
DB |||||  
QY 1021 GGCCTGTGAAAGGGAATCCATGCCATATGAGTTTGCATCAGGGTCCCGTTCTACGTG 1080  
DB |||||  
QY 1197 GGCCTGTGAAAGGGAATCCATGCCATATGAGTTTGCATCAGGGTCCCGTTCTACGTG 1256  
DB |||||

1081 AGGGGCCCAACGTGAAGCCGCTGTCTGAATCCCAATCGTCTCAACATTGACCTG 1140  
1257 AGGGGCCCAACGTGAAGCCGCTGTCTGAATCCCAATCGTCTCAACATTGACCTG 1316  
1141 GCCCCACCAATCTGTGACATTGACGCTGTGACATACCTGCGGATATGACGGGAATCC 1200  
1317 GCCCCACCAATCTGTGACATTGACGCTGTGACATACCTGCGGATATGACGGGAATCC 1376  
1201 ATCTCAAGCTGTGACACCGGAGCGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1260  
1377 ATCTCAAGCTGTGACACCGGAGCGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1436  
1261 AGGGTCTGGGGGACTCTCTTCTTGGTGGAGAGAGGCAAGCTGCTACACAAGAGACAAAT 1320  
1437 AGGGTCTGGGGGACTCTCTTCTTGGTGGAGAGAGGCAAGCTGCTACACAAGAGACAAAT 1496  
1321 GACAAGTGTGACCGCCACGAGGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGACCTG 1380  
1497 GACAAGTGTGACCGCCACGAGGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGACCTG 1556  
1381 TGTCAAGCTGTGAGTACAGACCGCGTGTGAGCAAGCTGGGACAGAGTGGCAGTGTGTG 1440  
1557 TGTCAAGCTGTGAGTACAGACCGCGTGTGAGCAAGCTGGGACAGAGTGGCAGTGTGTG 1616  
1441 GAGCAGCGCCACGGGGAACTGAAGCTGTGATAGTGCAGGGGCCCATGCGGCTGGCGGC 1500  
1617 GAGCAGCGCCACGGGGAACTGAAGCTGTGATAGTGCAGGGGCCCATGCGGCTGGCGGC 1676  
1501 AGCAGAGCCCTCTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGAGCCCTGCACC 1560  
1677 AGCAGAGCCCTCTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGAGCCCTGCACC 1736  
1561 TGTGACAGGGGGACTACAGCTCAGCTGCGGCGGACGCGGCAAAACTCTTCAAGAG 1620  
1737 TGTGACAGGGGGACTACAGCTCAGCTGCGGCGGACGCGGCAAAACTCTTCAAGAG 1796  
1621 AAGTACAAGGCCAGCTATGTCCGCGAGTCCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1680  
1797 AAGTACAAGGCCAGCTATGTCCGCGAGTCCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1856  
1681 GGCAGGGTGTACAGCTAGGCTGTGGTGTATGGCGCCAGCCCGGAAACTCTACAGCGG 1740  
1857 GGCAGGGTGTACAGCTAGGCTGTGGTGTATGGCGCCAGCCCGGAAACTCTACAGCGG 1916  
1741 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGGATGTGGGACTTTCAGTGGCACT 1800  
1917 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGGATGTGGGACTTTCAGTGGCACT 1976  
1801 GGAGGCTTCCCGACTACTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTTACATC 1860  
1977 GGAGGCTTCCCGACTACTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTTACATC 2036  
1861 CTAGAGAACGACACAGTCCAGTGTGACTGTGACCTGTATCAGTCCCTCAGGCTTGGAA 1920  
2037 CTAGAGAACGACACAGTCCAGTGTGACTGTGACCTGTATCAGTCCCTCAGGCTTGGAA 2096  
1921 GACCACAAGCTGCATCGACACAGAGATTGAACCCCTGCAGAACAAATTAAGAACCTG 1980  
2097 GACCACAAGCTGCATCGACACAGAGATTGAACCCCTGCAGAACAAATTAAGAACCTG 2156  
1981 AGGAGTCCGAGGTCACTTGAAGAAAAGCGGACAGAGATGTGACTGTACAAAATC 2040  
2157 AGGAGTCCGAGGTCACTTGAAGAAAAGCGGACAGAGATGTGACTGTACAAAATC 2216  
2041 AGCTACCAACCCAGCAAAAGCGGCTTCAAGCACAGAGCTCCAGTCTGCATCTTTC 2100  
2217 AGCTACCAACCCAGCAAAAGCGGCTTCAAGCACAGAGCTCCAGTCTGCATCTTTC 2276  
2101 AGAAGGGCTTGCAGAGAGGACAAAGTGTGGCTGTGGGAGCAGAGCGCAAGAG 2160  
2277 AGAAGGGCTTGCAGAGAGGACAAAGTGTGGCTGTGGGAGCAGAGCGCAAGAG 2336

2161 AAACTCCGCAAGCTGTCTAAGCGCTTCGAGAACACAGCACAGTGCAGCATGCCAGGCCTC 2220  
2337 AAACTCCGCAAGCTGTCTAAGCGCTTCGAGAACACAGCACAGTGCAGCATGCCAGGCCTC 2396  
2221 ACGTGTCTACCCACGACACAGGACTGGCAGACGGCGCTTTCTGGACACTGGGGCCT 2280  
2397 ACGTGTCTACCCACGACACAGGACTGGCAGACGGCGCTTTCTGGACACTGGGGCCT 2456  
2281 TTCTGTGCTTGCACACGCGCAACAAATTAACAGTACTGTGTGATGAGGACCAATCAATGAG 2340  
2457 TTCTGTGCTTGCACACGCGCAACAAATTAACAGTACTGTGTGATGAGGACCAATCAATGAG 2516  
2341 ACTCACAATTTCTCTTCTGTGATTTGCAATGTGGTCTCTAGAGTACTTGTATCTCAAC 2400  
2517 ACTCACAATTTCTCTTCTGTGATTTGCAATGTGGTCTCTAGAGTACTTGTATCTCAAC 2576  
2401 ACAGACCCCTACAGCTGATGAATGACGTGAACACACTGGACAGAGGATGTCTCAACCAAG 2460  
2577 ACAGACCCCTACAGCTGATGAATGACGTGAACACACTGGACAGAGGATGTCTCAACCAAG 2636  
2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520  
2637 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2696  
2521 ACTCGAAAACATGACCTGGGACTTAAAGATGAGGAGGCTATGAGCAATACAGGCAAGTTT 2580  
2697 ACTCGAAAACATGACCTGGGACTTAAAGATGAGGAGGCTATGAGCAATACAGGCAAGTTT 2756  
2581 CAGCGTCGAAAGTGGCCAGAAAATGAAGAGACCTTTCTTCAAAATCACTGGGACAACTGTGG 2640  
2757 CAGCGTCGAAAGTGGCCAGAAAATGAAGAGACCTTTCTTCAAAATCACTGGGACAACTGTGG 2816  
2641 GAAGCTGGGAAGTTTAAAGAACACAGAGGTGGACCTTCCAAAACATAGAGGATCACC 2700  
2817 GAAGCTGGGAAGTTTAAAGAACACAGAGGTGGACCTTCCAAAACATAGAGGATCACC 2876  
2701 TGACTGCACAGGCAATGAAGAACACATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760  
2877 TGACTGCACAGGCAATGAAGAACACATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2936  
2761 AGGAGCCTTGAGAAAGCAAGCAGCCTCTCTAGTCAAATGACAGATTTCTGGAGGATAAC 2820  
2937 AGGAGCCTTGAGAAAGCAAGCAGCCTCTCTAGTCAAATGACAGATTTCTGGAGGATAAC 2996  
2821 CAGCAGAGCAGAGATAACTTCAAGGAAGTCCATTTTGGCCCTGTCTTTGTTGATTA 2880  
2997 CAGCAGAGCAGAGATAACTTCAAGGAAGTCCATTTTGGCCCTGTCTTTGTTGATTA 3056  
2881 TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAAGTCACCACTAACCCCTCCC 2940  
3057 TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAAGTCACCACTAACCCCTCCC 3116  
2941 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000  
3117 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3176  
3001 CCCAAGGGGAAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTAAA 3060  
3177 CCCAAGGGGAAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTAAA 3236  
3061 TCCTCTTATTTCTTTGGTTGTCAAAAGAGAACTAAGAGAGGAGGACAGAGGCAAGCT 3120  
3237 TCCTCTTATTTCTTTGGTTGTCAAAAGAGAACTAAGAGAGGAGGACAGAGGCAAGCT 3296  
3121 GGAGAGGCTGAAAACAGTGCAGAGCTTTTGAATAGTCAAGTACAGCAAAAGAGATGA 3180  
3297 GGAGAGGCTGAAAACAGTGCAGAGCTTTTGAATAGTCAAGTACAGCAAAAGAGATGA 3356  
3181 CATTTACCTAGCACTATAAAACCTTGGTTCCTCTGAGGAACTGCTTCAATGTATATAT 3240  
3357 CATTTACCTAGCACTATAAAACCTTGGTTCCTCTGAGGAACTGCTTCAATGTATATAT 3416  
3241 GTGACTATTATTACATGTAATCAACATGGGAACTTTTAGGGGAACCTTAATAAGAAAATCCCAA 3300

Db 3417 GTGACTATTTCATGTAATCAACATGGGAACCTTTAGGGGAACCTAATAAGAAATCCCAA 3476  
QY 3301 TTTTCAGGAGTGGTGTGTAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3352  
Db 3477 TTTTCAGGAGTGGTGTGTAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3528  
RESULT 4  
US-09-970-287-11  
; Sequence 11, Application US/09970287  
; Publication No. US20030166283A1  
; GENERAL INFORMATION:  
; APPLICANT: GLUCKSMANN, Maria A.  
; APPLICANT: MEYERS, Rachel  
; APPLICANT: KAPPELLER-LIBERMANN, Rosana  
; APPLICANT: SILOS-SANTIAGO, Inmaculada  
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR  
; FILE REFERENCE: 10147-61U1  
; CURRENT APPLICATION NUMBER: US/09/970,287  
; CURRENT FILING DATE: 2001-10-03  
; PRIOR APPLICATION NUMBER: US 60/257,082  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 4397  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-970-287-11  
Query Match 99.5%; Score 3347.2; DB 13; Length 4397;  
Best Local Similarity 99.9%; Pred No. 0;  
Matches 3349; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GACTCCCGATCCCAAGAGAGACACAGATCAGCAAAAGAGATGGGCCCGGAGC 60  
Db 286 GACTCCCGATCCCAAGAGAGACACAGATCAGCAAAAGAGATGGGCCCGGAGC 345  
QY 61 CTCGCTGTGCTGCTCTCGCACTGTGTCTCCCTGCTGGGTGAGAGTGGCCCTTC 120  
Db 346 CTCGCTGTGCTGCTCTCGCACTGTGTCTCCCTGCTGGGTGAGAGTGGCCCTTC 405  
QY 121 CTGTCGCACACCGCTGAAAGGAGAGTTTCAGAGGAGCCGACAGAAACATCCGCCCAAC 180  
Db 406 CTGTCGCACACCGCTGAAAGGAGAGTTTCAGAGGAGCCGACAGAAACATCCGCCCAAC 465  
QY 181 ATCATCTGTGCTGCTGACGAGACGACAGGATGTGAGTGGGTTCATGCAAGTGAAC 240  
Db 466 ATCATCTGTGCTGCTGACGAGACGACAGGATGTGAGTGGGTTCATGCAAGTGAAC 525  
QY 241 AAGACCCGGCGCATCATGGAGCAGGGCGGGCGCACTTCATCAACGCTTCGTGACCA 300  
Db 526 AAGACCCGGCGCATCATGGAGCAGGGCGGGCGCACTTCATCAACGCTTCGTGACCA 585  
QY 301 CCATGTGTCCTCAGCTCTCCATCTCTCACTGCAAGTACGTCCCAACCAAC 360  
Db 586 CCATGTGTCCTCAGCTCTCCATCTCTCACTGCAAGTACGTCCCAACCAAC 645  
QY 361 ACCTACCAACATGAGAGATGCTCTCGCCCTCTGCGAGGACACAGCAGAGCGGC 420  
Db 646 ACCTACCAACATGAGAGATGCTCTCGCCCTCTGCGAGGACACAGCAGAGCGGC 705  
QY 421 ACCTTTCGGGTGACTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT 480  
Db 706 ACCTTTCGGGTGACTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT 765  
QY 481 AATGATCAACAGGCTCTAGTGCACCGCTGGAAGAGTGGGTGCACTCTTAA 540  
Db 766 AATGATCAACAGGCTCTAGTGCACCGCTGGAAGAGTGGGTGCACTCTTAA 825  
QY 541 AACTCCCGCTTTATAACTACAGCTGTGTGCGAAGCGGGTGAAGAAAGACGCGCTCC 600

Db 826 AACTCCCGCTTTTATAACTACAGCTGTGTGCGAAGCGGGTGAAGAGAACGACGCTCC 885  
QY 601 GACTACTCCAGGATTTACCTCAGAGACCTCATCAAAATGACAGCGTGTCTTCCGC 660  
Db 886 GACTACTCCAGGATTTACCTCAGAGACCTCATCAAAATGACAGCGTGTCTTCCGC 945  
QY 661 ACCTCAAGAAGATGTACCCGACAGGCGGAGTCTCTCATGCTCATCAGCCATCAGGCCCC 720  
Db 946 ACCTCAAGAAGATGTACCCGACAGGCGGAGTCTCTCATGCTCATCAGCCATCAGGCCCC 1005  
QY 721 CAGGCGCTGAGGATTTACGCCCCACAAATTTACGGCTCTTCCGAAACGATCTCAGCAC 780  
Db 1006 CAGGCGCTGAGGATTTACGCCCCACAAATTTACGGCTCTTCCGAAACGATCTCAGCAC 1065  
QY 781 ATCAGCGCGAGCTACAACTACGCGCCCAACCCGCAACAACTGATCATGCGCTACAGC 840  
Db 1066 ATCAGCGCGAGCTACAACTACGCGCCCAACCCGCAACAACTGATCATGCGCTACAGC 1125  
QY 841 GGGCCCATGAGCCCATCCATGGAATTTACAAATGCTCCAGCGGAAGCGCTTGCA 900  
Db 1126 GGGCCCATGAGCCCATCCATGGAATTTACAAATGCTCCAGCGGAAGCGCTTGCA 1185  
QY 901 ACCCTCATGTGCTGAGACCTCCATGAGAGAGATTTCAACATGCTGTTGAGAGCGGC 960  
Db 1186 ACCCTCATGTGCTGAGACCTCCATGAGAGAGATTTCAACATGCTGTTGAGAGCGGC 1245  
QY 961 GAGTGGACAAACGCTACATCGTATACACCGCCGACCAAGTTACCAATCGGCCAGTTT 1020  
Db 1246 GAGTGGACAAACGCTACATCGTATACACCGCCGACCAAGTTACCAATCGGCCAGTTT 1305  
QY 1021 GGCCTGTTGAAAGGAAATCCATGCTATGATGTTGATCATCAGGCTCCGTTCTACGTC 1080  
Db 1306 GGCCTGTTGAAAGGAAATCCATGCTATGATGTTGATCATCAGGCTCCGTTCTACGTC 1365  
QY 1081 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCTCAACATTCACCTG 1140  
Db 1366 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCTCAACATTCACCTG 1425  
QY 1141 GCGCCACCATCTCTGACATTCGAGCGCTGAGACATACCTGCGGATATGACGCGGAAATCC 1200  
Db 1426 GCGCCACCATCTCTGAGCATTTGACGCGCTGGACATACCTGCGGATATGACGCGGAAATCC 1485  
QY 1201 ATCTCTCAAGCTCTGACACAGGAGCGGCGGTTGATCGGTTTCACTTGAAAAGAGATG 1260  
Db 1486 ATCTCTCAAGCTCTGACACAGGAGCGGCGGTTGATCGGTTTCACTTGAAAAGAGATG 1545  
QY 1261 AGGCTCTGCGGGAATCTCTTTGTTGAGAGAGGCAAGTCTCTACAAAGAGAGACAAT 1320  
Db 1546 AGGCTCTGCGGGAATCTCTTTGTTGAGAGAGGCAAGTCTCTACAAAGAGAGACAAT 1605  
QY 1321 GACAGGTGAGCGCCGAGGAGAGAACTTCTGCGCAAGTACCGGCTGTGAGGACCTG 1380  
Db 1606 GACAGGTGAGCGCCGAGGAGAGAACTTCTGCGCAAGTACCGGCTGTGAGGACCTG 1665  
QY 1381 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGAGCTGGGACAGAACTGGCAGTGTGTG 1440  
Db 1666 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGAGCTGGGACAGAACTGGCAGTGTGTG 1725  
QY 1441 GAGGAGCGCACCGGGAAGCTGAGCTGATAGTCAAGGCGCCCATCGGCTGGGCGGC 1500  
Db 1726 GAGGAGCGCACCGGGAAGCTGAGCTGATAGTCAAGGCGCCCATCGGCTGGGCGGC 1785  
QY 1501 AGCAGAGCGCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCAAC 1560  
Db 1786 AGCAGAGCGCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCAAC 1845  
QY 1561 TGTGACGCGGGAATCAAGCTAGCTGCGCGGACCGCGGAAAAAACTCTTCAAGAAG 1620  
Db 1846 TGTGACGCGGGAATCAAGCTAGCTGCGCGGACCGCGGAAAAAACTCTTCAAGAAG 1905  
QY 1621 AAGTACAAGGCCAGCTATGTCCGAGTGTGCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1680  
Db 1906 AAGTACAAGGCCAGCTATGTCCGAGTGTGCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1965

1681 GGCAGGGTGTACCACTAGGCTGGGTGATGCGCCAGCCCGGAAACCTCAACAGCGG 1740  
1666 GGCAGGGTGTACCACTAGGCTGGGTGATGCGCCAGCCCGGAAACCTCAACAGCGG 2025  
1741 CACTGGCAGGGGCCCTGTAGGACCAAGATGCAAGATGATGGGACCTTCAGTGGCACT 1800  
2026 CACTGGCAGGGGCCCTGTAGGACCAAGATGCAAGATGATGGGACCTTCAGTGGCACT 2085  
1801 GGAGGCTTCCCGACTACTACTACGCGCCCAACCCATTAAGTGTACACATCGGTGTACATC 1860  
2086 GGAAGGCTTCCCGACTACTACTACGCGCCCAACCCATTAAGTGTACACATCGGTGTACATC 2145  
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGACGCGCTGGAAA 1920  
2146 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGACGCGCTGGAAA 2205  
1921 GACCAACAGCTGCATCGACACAGAGTGAACCCCTGCGACACAAATTAAGAACCTG 1980  
2206 GACCAACAGCTGCATCGACACAGAGTGAACCCCTGCGACACAAATTAAGAACCTG 2265  
1981 AGGGAAGTCCGAGGTCACTCTGAAGAAAAGCGGCCAGAGAAATGTGACTGTCAAAAATC 2040  
2266 AGGGAAGTCCGAGGTCACTCTGAAGAAAAGCGGCCAGAGAAATGTGACTGTCAAAAATC 2325  
2041 AGCTACCAACCCAGGACAAAGCGCCCTCAAGCACAGAGGTCAGTCTGCATCTCTTTC 2100  
2326 AGCTACCAACCCAGGACAAAGCGCCCTCAAGCACAGAGGTCAGTCTGCATCTCTTTC 2385  
2101 AGGAAGGGCTCCAGAGAGAGGACAAAGTGTGGCTGTGTTGCGGAGCAGAAAGCAAGAG 2160  
2386 AGGAAGGGCTCCAGAGAGAGGACAAAGTGTGGCTGTGTTGCGGAGCAGAAAGCAAGAG 2445  
2161 AAATCCGCAAGTGTCTAAGCGCTCGAGAACAAACGACAGTGCAGATGCCAGGCTC 2220  
2446 AAATCCGCAAGTGTCTAAGCGCTCGAGAACAAACGACAGTGCAGATGCCAGGCTC 2505  
2221 AGTGTCTTCAACCCAGACACAGCACTGGCAGAGCGGCTTTCTGACACATGGGCGCT 2280  
2506 AGTGTCTTCAACCCAGACACAGCACTGGCAGAGCGGCTTTCTGACACATGGGCGCT 2565  
2281 TTCTGTGCTGCAACGCGCCCAACAAATACAGTACTGTGTGATGAGGACCAATCAATGAG 2340  
2566 TTCTGTGCTGCAACGCGCCCAACAAATACAGTACTGTGTGATGAGGACCAATCAATGAG 2625  
2341 ACTCAAAATTTCTCTCTGTAATTTGCAATGCTGCTTCTAGAGTACTTTGATCTCAAC 2400  
2626 ACTCAAAATTTCTCTCTGTAATTTGCAATGCTGCTTCTAGAGTACTTTGATCTCAAC 2685  
2401 ACAGACCCCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
2686 ACAGACCCCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2745  
2461 CTACAGTACAGTCTATGAGTCTGAGGAGCTGCAAGGGTTACAGCAGTGTAAACCCCGG 2520  
2746 CTACAGTACAGTCTATGAGTCTGAGGAGCTGCAAGGGTTACAGCAGTGTAAACCCCGG 2805  
2521 ACTCGAAACATGCACTGGGACTTAAAGATGAGGAGCTATGAGCAATACAGGCGAGTTT 2580  
2806 ACTCGAAACATGCACTGGGACTTAAAGATGAGGAGCTATGAGCAATACAGGCGAGTTT 2865  
2581 CAGCGTGAAGTGGCCAGAAATGAAGAGACCTTTTCCAAATCACTGGGCAACTGTGG 2640  
2866 CAGCGTGAAGTGGCCAGAAATGAAGAGACCTTTTCCAAATCACTGGGCAACTGTGG 2925  
2641 GAAGGCTGGGAGGTTTAAAGAAACACAGAGGTGGACCTTCCAAAACATAGAGGATCACC 2700  
2926 GAAGGCTGGGAGGTTTAAAGAAACACAGAGGTGGACCTTCCAAAACATAGAGGATCACC 2985  
2701 TGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 2760  
2986 TGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 3045

2761 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGAGGATAAC 2820  
3046 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGAGGATAAC 3105  
2821 CAGCAGGACGAGATTAATCTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTGA 2880  
3106 CAGCAGGACGAGATTAATCTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTGA 3165  
2881 TACTCTCAGCAGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCAACCACTAACCTCC 2940  
3166 TACTCTCAGCAGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCAACCACTAACCTCC 3225  
2941 CCAGAAAGCTCAAAAAGGAAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000  
3226 CCAGAAAGCTCAAAAAGGAAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3285  
3001 CCCAAGGGGAAAGTCAATGGAAATTTTAAATCATAGGGGAAAACAGTCTGTTCTAAA 3060  
3286 CCCAAGGGGAAAGTCAATGGAAATTTTAAATCATAGGGGAAAACAGTCTGTTCTAAA 3345  
3061 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGAACTAAGAAAGCAGGACAGGCAACGT 3120  
3346 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGAACTAAGAAAGCAGGACAGGCAACGT 3405  
3121 GAGAGGCTGAAAACAGTGCAGAGAGCTTTGACATGAGTCACTAGGACAAAAGAGATGA 3180  
3406 GAGAGGCTGAAAACAGTGCAGAGAGCTTTGACATGAGTCACTAGGACAAAAGAGATGA 3465  
3181 CATTTACCTAGCACTATAAACCTGTTGCTCTCTGAAGAACTGCTTCAATGTATATAT 3240  
3466 CATTTACCTAGCACTATAAACCTGTTGCTCTCTGAAGAACTGCTTCAATGTATATAT 3525  
3241 GTGACTATTTACATGATTAATCAACATGGGAACTTTTAGGGAACTTAATAGAAATCCCAA 3300  
3526 GTGACTATTTACATGATTAATCAACATGGGAACTTTTAGGGAACTTAATAGAAATCCCAA 3585  
3301 TTTTCAGGAGTGGTGGTGTCAATAAAGCTCTGTGCGCAGTGTAAAAAGAAA 3352  
3586 TTTTCAGGAGTGGTGGTGTCAATAAAGCTCTGTGCGCAGTGTAAAAAGAAA 3637

## RESULT 5

US-10-025-966A-13  
; Sequence 13, Application US/10025966A  
; Publication No. US20030148920A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Steven D.  
; APPLICANT: Palmeri, Diana  
; TITLE OF INVENTION: Sulfatases and methods of use  
; TITLE OF INVENTION: thereof  
; FILE REFERENCE: UCAL230  
; CURRENT APPLICATION NUMBER: US/10/025,966A  
; CURRENT FILING DATE: 2002-12-21  
; PRIOR APPLICATION NUMBER: 60/258,577  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 60/267,831  
; PRIOR FILING DATE: 2001-09-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 4286  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-025-966A-13

Query Match 99.5%; Score 3346.8; DB 13; Length 4286;  
Best Local Similarity 99.9%; Pred.No. 0;  
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GACTCCCGCATCCCAAAGAGCAGATCAGCAAAAAGAGATGGGCCCGGAGC 60  
DB 632 GACTCCCGCATCCCAAAGAGCAGATCAGCAAAAAGAGATGGGCCCGGAGC 691









QY 541 AACTCCGCTTTTATACACACGCTGTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCC 600  
DB 1172 AACTCCGCTTTTATACACACGCTGTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCC 1231  
QY 601 GACTACTCCAAAGATTACCTACAGACCTCATACCAATGACAGCGTGTCTTCGCG 660  
DB 1232 GACTACTCCAAAGATTACCTACAGACCTCATACCAATGACAGCGTGTCTTCGCG 1291  
QY 661 AGCTCAAGAAAGATTGATCCCGCACAGGCGAGTCTCATGGTCAATGACAGCGTGTCTTCGCG 720  
DB 1292 AGCTCAAGAAAGATTGATCCCGCACAGGCGAGTCTCATGGTCAATGACAGCGTGTCTTCGCG 1351  
QY 721 CACGCGCTGTAGGATTACGCCCCACCAATATTACGCGCTCTCCCAAGCGATCTCAGAC 780  
DB 1352 CACGCGCTGTAGGATTACGCCCCACCAATATTACGCGCTCTCTCCCAAGCGATCTCAGAC 1411  
QY 781 ATACGCGCGAGTCAACACTACGCGCGCCCAACCGGACAAACACTGATCATCGCTTACACG 840  
DB 1412 ATACGCGCGAGTCAACACTACGCGCGCCCAACCGGACAAACACTGATCATCGCTTACACG 1471  
QY 841 GGGCCCATGAGCCCATCCACATGAAATTCACCAACATGCTCCAGCGGAAGCGTTCGAG 900  
DB 1472 GGGCCCATGAGCCCATCCACATGAAATTCACCAACATGCTCCAGCGGAAGCGTTCGAG 1531  
QY 901 ACCCTCATGTCCGTGGAGCGACTCCATGAGACGATTTACAACTGCTGTTGACACGGC 960  
DB 1532 ACCCTCATGTCCGTGGAGCGACTCCATGAGACGATTTACAACTGCTGTTGACACGGC 1591  
QY 961 GAGCTGGACACACATGATACATCGTATACACCGCGACACCGGTTACCAATCGCGCCAGTTT 1020  
DB 1592 GAGCTGGACACACATGATACATCGTATACACCGCGACACCGGTTACCAATCGCGCCAGTTT 1080  
QY 1021 GGCCTGGTGAAGGAAATCCATGCTCATATGAGTTGACATCAGGCTCCGCTTCTACGTG 1080  
DB 1652 GGCCTGGTGAAGGAAATCCATGCTCATATGAGTTGACATCAGGCTCCGCTTCTACGTG 1711  
QY 1081 AGGGCCCCAAGCTGGAAAGCGGCTGTCTGAATCCCACTCGTCTCAACATGACCTG 1140  
DB 1712 AGGGCCCCAAGCTGGAAAGCGGCTGTCTGAATCCCACTCGTCTCAACATGACCTG 1771  
QY 1141 GCGCCCATGCTGGACATTTGACGCGCTGACATACCTGCGGATGTGAGCGGAATCC 1200  
DB 1772 GCGCCCATGCTGGACATTTGACGCGCTGACATACCTGCGGATGTGAGCGGAATCC 1831  
QY 1201 ATCTCTCAAGCTGCTGGACACGAGCGGCGGCTGAATCGTTTCACTTTGAAAAGAGATG 1260  
DB 1832 ATCTCTCAAGCTGCTGGACACGAGCGGCGGCTGAATCGTTTCACTTTGAAAAGAGATG 1891  
QY 1261 AGGCTCTGGCGGACTCTCTTCTGGTGAGAGAGCGAAGCTGTCTPACAAAGAGACAAT 1320  
DB 1892 AGGCTCTGGCGGACTCTCTTCTGGTGAGAGAGCGAAGCTGTCTPACAAAGAGACAAT 1951  
QY 1321 GACAGGTGGACGCGCCAGGAGGAACTTTCTGCGCCCAAGTACACAGCGTGTGAGGACCTG 1380  
DB 1952 GACAGGTGGAGCGCCAGGAGGAACTTTCTGCGCCCAAGTACACAGCGTGTGAGGACCTG 2011  
QY 1381 TGTCAAGCTGTGAGTACACAGCGCGCTGTGAGAGCTGGGACAGAAAGTGGCAGTGTG 1440  
DB 2012 TGTCAAGCTGTGAGTACACAGCGCGCTGTGAGAGCTGGGACAGAAAGTGGCAGTGTG 2071  
QY 1441 GAGAGCGCCACGGGAGCTGAAGCTGTGATAGTGAAGGCGCCCATGCGGCTGGCGCGC 1500  
DB 2072 GAGAGCGCCACGGGAGCTGAAGCTGTGATAGTGAAGGCGCCCATGCGGCTGGCGCGC 2131  
QY 1501 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGCGGAGGCGAGCGAGGCTGCAAC 1560  
DB 2132 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGCGGAGGCGAGCGAGGCTGCAAC 2191  
QY 1561 TGTGACAGCGGAGCTACAGAGCTCAGGCTGGCGGACGCGGAAAAAATCTTCAAGAG 1620  
DB 2192 TGTGACAGCGGAGCTACAGAGCTCAGGCTGGCGGAGCGCGGAAAAAATCTTCAAGAG 2251  
QY 1621 AAGTACAGGCGAGCTATGTCCGCGAGTCTCCATCCGCTCAGTGGGCCATCGAGGTGGAC 1680

DB 2252 AAGTACAGGCGAGCTATGTCCGAGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2311  
QY 1681 GGCAGGGTGTACACAGTACAGCTGGCTGGTGTATGCGGCCACCGCCGAAACCTCACCAGCGG 1740  
DB 2312 GGCAGGGTGTACACAGTACAGCTGGCTGGTGTATGCGGCCACCGCCGAAACCTCACCAGCGG 2371  
QY 1741 CACTGGCCCAAGGGGCCCTCAGGACCAAGATGACAAAGATGGTGGGACTTTCAGTGGCACT 1800  
DB 2372 CACTGGCCCAAGGGGCCCTCAGGACCAAGATGACAAAGATGGTGGGACTTTCAGTGGCACT 2431  
QY 1801 GAGGGCTTCCGAGTACTACTCAGCGGCAACCCCAATTAAGTGCACATCGGTGTACATC 1860  
DB 2432 GAGGGCTTCCGAGTACTACTCAGCGGCAACCCCAATTAAGTGCACATCGGTGTACATC 2491  
QY 1861 CTAGAGAAACACACAGTGTCAAGTGTGACCTGGACCTGTCAAGTCCCTGAGGCTTGAAA 1920  
DB 2492 CTAGAGAAACACACAGTGTCAAGTGTGACCTGGACCTGTCAAGTCCCTGAGGCTTGAAA 2551  
QY 1921 GACCAACAGCTGCACATCCACACAGAGTTGAAACCTCTGAGAACAAAATTAAGAACCTG 1980  
DB 2552 GACCAACAGCTGCACATCCACACAGAGATTGAAACCTCTGAGAACAAAATTAAGAACCTG 2611  
QY 1981 AGGGAAGTCCGAGGTCACCTGAAGAAAGCGGCCAGAGAGATGTGACTGTCACAAAATC 2040  
DB 2612 AGGGAAGTCCGAGGTCACCTGAAGAAAGCGGCCAGAGAGATGTGACTGTCACAAAATC 2671  
QY 2041 AGCTACCAACCCAGCACAAAAGCCGCTCAAGACACAGAGGCTCCAGTCTGCTCTTTC 2100  
DB 2672 AGCTACCAACCCAGCACAAAAGCCGCTCAAGACACAGAGGCTCCAGTCTGCTCTTTC 2731  
QY 2101 AGGAGGGCTCCAGAGAGAGACAAAGTGTGCTGTTGCGGAGACAGAGCGCAAGAG 2160  
DB 2732 AGGAGGGCTCCAGAGAGAGACAAAGTGTGCTGTTGCGGAGACAGAGCGCAAGAG 2791  
QY 2161 AAATCCGCAAGCTGTCAAGCGCTGAGAAACACAGACACAGTGCAGATGCCAGGCTC 2220  
DB 2792 AAATCCGCAAGCTGTCAAGCGCTGAGAAACACAGACACAGTGCAGATGCCAGGCTC 2851  
QY 2221 AGTGTCTTCCACCAAGCAACAGCACTGGCAGACGGCGCTTCTGCACTCTGGGGCT 2280  
DB 2852 AGTGTCTTCCACCAAGCAACAGCACTGGCAGACGGCGCTTCTGCACTCTGGGGCT 2911  
QY 2281 TTCTGTGCTGCACAGCGCCAAACATTAACGTAAGTGTGATGAGGACCAATCAATGAG 2340  
DB 2912 TTCTGTGCTGCACAGCGCCAAACATTAACGTAAGTGTGATGAGGACCAATCAATGAG 2971  
QY 2341 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400  
DB 2972 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 3031  
QY 2401 ACAGACCCCTACAGCTGATGATGACAGTGAACACACTGGACAGGATGTCTCAACAG 2460  
DB 3032 ACAGACCCCTACAGCTGATGATGACAGTGAACACACTGGACAGGATGTCTCAACAG 3091  
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTACAGAGCTGTAAACCCCGG 2520  
DB 3092 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTACAGAGCTGTAAACCCCGG 3151  
QY 2521 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2580  
DB 3152 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 3211  
QY 2581 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640  
DB 3212 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 3271  
QY 2641 GAAGCTGGGAGGTTAAAGAAACAAACAGAGGTGGACCTTCCAAAAACATAGAGGATCAAC 2700  
DB 3272 GAAGCTGGGAGGTTAAAGAAACAAACAGAGGTGGACCTTCCAAAAACATAGAGGATCAAC 3331  
QY 2701 TGACTGCACAGGCAATGAAGAACCAATGAGGAGTGTGCTTCCAGACAGCTGTCTATTGGCC 2760

```
Db 3332 TGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCACAGACCTGTGTCTATTGGCC 3391
Qy 2761 AGGAGGCTCTGAGAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3392 AGGAGGCTCTGAGAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3451
Qy 2821 CAGCAGAGCAGAGATTAATCTTCAGGAAGTCATTTTGGCCCTGCTTTTGGCTTTGGATTA 2880
Db 3452 CAGCAGAGCAGAGATTAATCTTCAGGAAGTCATTTTGGCCCTGCTTTTGGCTTTGGATTA 3511
Qy 2881 TACCTCACCAAGCTGCACAAATGCAATTTTGTATCAAAAGTCAACCTAACCCCTCCC 2940
Db 3512 TACCTCACCAAGCTGCACAAATGCAATTTTGTATCAAAAGTCAACCTAACCCCTCCC 3571
Qy 2941 CCAGAACTCACAAGGAAACCGAGAGCGAGCGAGAGATTTCTTGGAAATTTCT 3000
Db 3572 CCAGAACTCACAAGGAAACCGAGAGCGAGCGAGAGATTTCTTGGAAATTTCT 3631
Qy 3001 CCCAAGGGCGAAAGTCAATTCGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAA 3060
Db 3632 CCCAAGGGCGAAAGTCAATTCGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAA 3691
Qy 3061 TCCTCTTATTTCTTGGTTTGTCAAAAGAACTTAAGAACGAGGCAACCT 3120
Db 3692 TCCTCTTATTTCTTGGTTTGTCAAAAGAACTTAAGAACGAGGCAACCT 3751
Qy 3121 GGAGAGGCTCAAAACAGTCAGAGAGCGTTTGACAACTAGTACAGTACGACAAAGAGATGA 3180
Db 3752 GGAGAGGCTCAAAACAGTCAGAGAGCGTTTGACAACTAGTACAGTACGACAAAGAGATGA 3811
Qy 3181 CATTTACCTAGCACTATAACCTCTGCTCTGAAGAACTGCTTCAATTTGTATATAT 3240
Db 3812 CATTTACCTAGCA-TATAACCTCTGCTCTGAAGAACTGCTTCAATTTGTATATAT 3870
Qy 3241 GTGACTATTTACATGTATCAACATGGGAACCTTTAGGGGAACCTTAAGAAATCCCAA 3300
Db 3871 GTGACTATTTACATGTATCAACATGGGAACCTTTAGGGGAACCTTAAGAAATCCCAA 3390
Qy 3301 TTTTCAGGAGTGTGTGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAGAAAGAA 3360
Db 3391 TTTTCAGGAGTGTGTGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAGAAAGAA 3990
Qy 3361 AA 3362
Db 3991 AA 3992

RESULT 7
; Sequence 17, Application US/10182951
; Publication No. US20030138895A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameera R.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SHIH, Leo L.
; APPLICANT: YANG, Junming
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: FI-0033 PCT
; CURRENT APPLICATION NUMBER: US/10/182,951
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/181,856; 10/183,684; 60/185,141; 60/186,818; 60/188,345;
; PRIOR FILING DATE: 2000-02-11; 2000-02-17; 2000-02-25; 2000-03-03; 2000-03-09; 2000-0
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 17
```

```
; LENGTH: 3348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030138895A1 1558210CB1
US-10-182-951-17

Query Match          99.5%; Score 3344.8; DB 13; Length 3348;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 12  CCCAAAAGAACACACAGATCAGCAAAAAGAGATGGCCCCCGAGCCTCGTGTGTG 71
Db 1  CCCAAAAGAACACACAGATCAGCAAAAAGAGATGGCCCCCGAGCCTCGTGTGTG 60
Qy 72  CTTGCTGTCCGCACTGTCTCTCTGCTGGTGGAGTCTGGCCTTCTGTGCAACA 131
Db 61  CTTGCTGTCCGCACTGTCTCTCTGCTGGTGGAGTCTGGCCTTCTGTGCAACA 120
Qy 132 CCGCTCGAAAAGGCGAGGTTTCAGAGGGACCGCAGAACATCCGCCCAACATCATCTCTGT 191
Db 121 CCGCTCGAAAAGGCGAGGTTTCAGAGGGACCGCAGAACATCCGCCCAACATCATCTCTGT 180
Qy 192 GCTCAACGACGACACAGATGTGAGTGGTTCATGACAGTGAAGAACACCGGG 251
Db 181 GCTCAACGACGACACAGATGTGAGTGGTTCATGACAGTGAAGAACACCGGG 240
Qy 252 CATCATGGAGCAGGGCGGCGCACTTCATCAACGCTTCGTGACCAACACCCATGTGCTG 311
Db 241 CATTTAGAGCAGGGCGGCGCACTTCATCAACGCTTCGTGACCAACACCCATGTGCTG 300
Qy 312 CCCCTCAGGCTCTCTCATCTCTCACTGGCAAGTACGTCCACCAACACACCTTACACAA 371
Db 301 CCCCTCAGGCTCTCTCATCTCTCACTGGCAAGTACGTCCACCAACACACCTTACACAA 360
Qy 372 CAATGAGAACTGTCTCTCGCCCTCTGGAGGACACAGCAGAGCGGCGACCTTTCGGCT 431
Db 361 CAATGAGAACTGTCTCTCGCCCTCTGGAGGACACAGCAGAGCGGCGACCTTTCGGCT 420
Qy 432 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATACAA 491
Db 421 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATACAA 480
Qy 492 CGGCTCCTAGTGGCCACCGGCTTGGAGAGTGGTTCGACTCTTAAATACTCCGCTT 551
Db 481 CGGCTCCTAGTGGCCACCGGCTTGGAGAGTGGTTCGACTCTTAAATACTCCGCTT 540
Qy 552 TTATAACTACACGCTGTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCCGACTACTCCAA 611
Db 541 TTATAACTACACGCTGTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCCGACTACTCCAA 600
Qy 612 GGATTACCTCAGACCTCATACCAATGACAGCGTGGTTCCTCCGACGTCGCAAGAA 671
Db 601 GGATTACCTCAGACCTCATACCAATGACAGCGTGGTTCCTCCGACGTCGCAAGAA 660
Qy 672 GATGTACCCGACAGGCGCAGTCTCTCATGTGTATGATCATGAGCGCCCGCCCGCTGTA 731
Db 661 GATGTACCCGACAGGCGCAGTCTCTCATGTGTATGATCATGAGCGCCCGCCCGCTGTA 720
Qy 732 GGATTACCCGACAGGCGCAGTCTCTCATGTGTATGATCATGAGCGCCCGCCCGCTGTA 791
Db 721 GGATTACCCGACAGGCGCAGTCTCTCATGTGTATGATCATGAGCGCCCGCCCGCTGTA 780
Qy 792 CTACAACCTAGCGCCCAACCCGAGCAACACATGATCATGCGCTACACGGGCGCCATGAA 851
Db 781 CTACAACCTAGCGCCCAACCCGAGCAACACATGATCATGCGCTACACGGGCGCCATGAA 840
Qy 852 GCCCATCCCATGGAATTCACCAACATGTCTCCAGCGGAAGCGCTTGCAGACCTCATGTC 911
Db 841 GCCCATCCCATGGAATTCACCAACATGTCTCCAGCGGAAGCGCTTGCAGACCTCATGTC 900
Qy 912 GGTGGAGCACTCCATGAGAGCGATTTTACAACATGTGTTGAGACGGGCGAGCTGACAA 971
```

Db 901 GGTGGACGACTCATGGAGACGATTTCACCATGCTGGTTGAGACGGCGAGCTGGACAA 960  
QY 972 CACGTACATCGTATACACCGCGACCAACCGTTTACACATCGGCCCACTTTGGGCTCGTGA 1031  
Db 961 CACGTACATCGTATACACCGCGACCAACCGTTTACACATCGGCCCACTTTGGGCTCGTGA 1020  
QY 1032 AGGGAATCCATGCGCATATAGTTTGACATCAGGGTCCGTTCTACTGAGGGGCCCCAA 1091  
Db 1021 AGGGAATCCATGCGCATATAGTTTGACATCAGGGTCCGTTCTACTGAGGGGCCCCAA 1080  
QY 1092 CGTGAAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTGACCTGGGCCCCACCAT 1151  
Db 1081 CGTGAAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTGACCTGGGCCCCACCAT 1140  
QY 1152 CTTGGACATTCAGAGGCTGGACATACCTTCGGATATGACGCGGAATCCATCTCAAGCT 1211  
Db 1141 CTTGGACATTCAGAGGCTGGACATACCTTCGGATATGACGCGGAATCCATCTCAAGCT 1200  
QY 1212 GCTGGACCGGAGCGCGGTGATCGGTTTCACTTGAAGAGAGATGAGGGTCTGGCG 1271  
Db 1201 GCTGGACCGGAGCGCGGTGATCGGTTTCACTTGAAGAGAGATGAGGGTCTGGCG 1260  
QY 1272 GGAATCCTTCTTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGA 1331  
Db 1261 GGAATCCTTCTTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGA 1320  
QY 1332 CGCCAGGAGGAGAACTTCTGCCCAAGTACACAGCGTGTGAAGGACCTGTGTGAGCGTGC 1391  
Db 1321 CGCCAGGAGGAGAACTTCTGCCCAAGTACACAGCGTGTGAAGGACCTGTGTGAGCGTGC 1380  
QY 1392 TGAGTACACAGCGGCTGTGACAGCTGGGACAGAAGTGGCAGTGTGTGGAGGAGCGCAC 1451  
Db 1381 TGAGTACACAGCGGCTGTGACAGCTGGGACAGAAGTGGCAGTGTGTGGAGGAGCGCAC 1440  
QY 1452 GGGGAAGCTGAAGCTGCATTAAGTGAAGGCGGCCCATCGCGCTGGCGCGCAGCAGAGCCCT 1511  
Db 1441 GGGGAAGCTGAAGCTGCATTAAGTGAAGGCGGCCCATCGCGCTGGCGCGCAGCAGAGCCCT 1500  
QY 1512 CTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGGAGCGCTGCACCTGTGACAGCGG 1571  
Db 1501 CTCCAACTCTGTGCCCAAGTACTACGGGACGGGACGGAGCGCTGCACCTGTGACAGCGG 1560  
QY 1572 GGACTACAAGCTCAGCTGGCGGACGGCGGAACTTCTCAAGAGAGAGTACAAGGC 1631  
Db 1561 GGACTACAAGCTCAGCTGGCGGACGGCGGAACTTCTCAAGAGAGAGTACAAGGC 1620  
QY 1632 CAGCTATGTCGACGCTGCATCGCTCAGTGGCCATCAGAGTGAAGCGGAGGCTGTA 1691  
Db 1621 CAGCTATGTCGACGCTGCATCGCTCAGTGGCCATCAGAGTGAAGCGGAGGCTGTA 1680  
QY 1692 CCAGTATGTCGCTGGTGTATGCGGCCAGCGCCGAACTCACCAGCGGCACTGGCCAGG 1751  
Db 1681 CCAGTATGTCGCTGGTGTATGCGGCCAGCGCCGAACTCACCAGCGGCACTGGCCAGG 1740  
QY 1752 GGCCTCTGAGGACCAAGATGACAAGATGTTGGGGAATTCAGTGGCACTGGAGCGCTTCC 1811  
Db 1741 GGCCTCTGAGGACCAAGATGACAAGATGTTGGGGAATTCAGTGGCACTGGAGCGCTTCC 1800  
QY 1812 CGACTACTCAGCGCGCAACCCATTAAAGTGAACATCGGTGTACATCTCTAGAGACGA 1871  
Db 1801 CGACTACTCAGCGCGCAACCCATTAAAGTGAACATCGGTGTACATCTCTAGAGACGA 1860  
QY 1872 CACGTCCAGTGTGACTGGACCTGTACAAGTCCCTGAGGCTGGAAAGACCAAGACT 1931  
Db 1861 CACGTCCAGTGTGACTGGACCTGTACAAGTCCCTGAGGCTGGAAAGACCAAGACT 1920  
QY 1932 GCACATCCAGCAGAGATTGAACCTCGAGAACAAATTAAGAACCTGAGGAGTCCG 1991  
Db 1921 GCACATCCAGCAGAGATTGAACCTCGAGAACAAATTAAGAACCTGAGGAGTCCG 1980  
QY 1992 AGGTCACTGAAGAAAAAGCGGCGCAGAGAGATGTGACTGTCAAAAATCAGCTACCCAC 2051

Db 1981 AGGTCACTGAAGAAAAAGCGCGCAGAGAAATGTGACTGTCAAAAATCAGCTACCCAC 2040  
QY 2052 CCAGCACAAGGCGGCTCAAGCACAAGAGCTCCAGTCTGCATCTCTTTCAGGAAGGCGCT 2111  
Db 2041 CCAGCACAAGGCGGCTCAAGCACAAGAGCTCCAGTCTGCATCTCTTTCAGGAAGGCGCT 2100  
QY 2112 GCAAGAGAGAGCAAGAGTGTGGCTGTGGGGAGCAGAGCGCAAGAGAACTCCGCGAA 2171  
Db 2101 GCAAGAGAGAGCAAGAGTGTGGCTGTGGGGAGCAGAGCGCAAGAGAACTCCGCGAA 2160  
QY 2172 GCTGCTCAAGCGCTCGAGAACACACACAGCTGCAGATGCCAGGCTCACTGCTTCAC 2231  
Db 2161 GCTGCTCAAGCGCTCGAGAACACACACAGCTGCAGATGCCAGGCTCACTGCTTCAC 2220  
QY 2232 CCAGCACAAGCAGCTGGCAGACGGCGCTTCTCGACACTGGGCGCTTCTGTGCTG 2291  
Db 2221 CCAGCACAAGCAGCTGGCAGACGGCGCTTCTCGACACTGGGCGCTTCTGTGCTG 2280  
QY 2292 CACAGGCGCAACAAATACAGTACTGTGTGCATGAGGACCAATGAGACTCAAAATTT 2351  
Db 2281 CACAGGCGCAACAAATACAGTACTGTGTGCATGAGGACCAATGAGACTCAAAATTT 2340  
QY 2352 CTTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTCTCAACACAGACCCCTA 2411  
Db 2341 CTTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTCTCAACACAGACCCCTA 2400  
QY 2412 CCACTGTATGAATGCAAGTGAACACACTGCAAGGGATGCTCAACCACTACAGTACA 2471  
Db 2401 CCACTGTATGAATGCAAGTGAACACACTGCAAGGGATGCTCAACCACTACAGTACA 2460  
QY 2472 GCTCATGGAGCTGAGGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGGACTCGAAACAT 2531  
Db 2461 GCTCATGGAGCTGAGGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGGACTCGAAACAT 2520  
QY 2532 GGAACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGTTCAGCGTCGAAA 2591  
Db 2521 GGAACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGTTCAGCGTCGAAA 2580  
QY 2592 GTGGCCAGAAATGAGAGAGCTTCTTCCAAATCACTGGGACAACTGTGGGAAAGGCTGGA 2651  
Db 2581 GTGGCCAGAAATGAGAGAGCTTCTTCCAAATCACTGGGACAACTGTGGGAAAGGCTGGA 2640  
QY 2652 AGGTTAAGAAACAAACAGAGGCTGGAGCTTCCAAAACATAGAGGATCACTGACTGCACAG 2711  
Db 2641 AGGTTAAGAAACAAACAGAGGCTGGAGCTTCCAAAACATAGAGGATCACTGACTGCACAG 2700  
QY 2712 GCATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCCAGAGGCTGA 2771  
Db 2701 GCATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCCAGAGGCTGA 2760  
QY 2772 GAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCCTGGAGGATTAACAGCAGAGCA 2831  
Db 2761 GAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCCTGGAGGATTAACAGCAGAGCA 2820  
QY 2832 GAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGTCTTTGCTTTGGATTATACCTCACCAG 2891  
Db 2821 GAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGTCTTTGCTTTGGATTATACCTCACCAG 2880  
QY 2892 CTGCAAAAATGCAATTTTCTGTATCAAAAAGTCAACCACTAACCCCTCCCGAGAGCTCA 2951  
Db 2881 CTGCAAAAATGCAATTTTCTGTATCAAAAAGTCAACCACTAACCCCTCCCGAGAGCTCA 2940  
QY 2952 CAAAGGAAACCGAGAGAGCGAGGAGAGATTTCTTTGGAAATTTCTCCCAAGGGCGA 3011  
Db 2941 CAAAGGAAACCGAGAGAGCGAGGAGAGATTTCTTTGGAAATTTCTCCCAAGGGCGA 3000  
QY 3012 AAGTCAATGGAAATTTTAAATCAATAGGGGAAAGCAGTCTGTCTTAAATCTCTTATTC 3071  
Db 3001 AAGTCAATGGAAATTTTAAATCAATAGGGGAAAGCAGTCTGTCTTAAATCTCTTATTC 3060  
QY 3072 TTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGAGAGGCAACCTGGAGAGGCTGA 3131  
Db 3061 TTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGAGAGGCAACCTGGAGAGGCTGA 3120







QY 1021 GGCCTGCTGAAGGGAATCCATGCCATATGAGTTTGAATCAAGGCTCCGCTTCTACGTTG 1080  
Db 1583 GGCCTGCTGAAGGGAATCCATGCCATATGAGTTTGAATCAAGGCTCCGCTTCTACGTTG 1642  
QY 1081 AGGGGCCCCAACGCTGGAAGCGCGTGTCTGAATCCCCACATCGTCTCAACATTGACCTG 1140  
Db 1643 AGGGGCCCCAACGCTGGAAGCGCGTGTCTGAATCCCCACATCGTCTCAACATTGACCTG 1702  
QY 1141 GCCCCCAACCATCTCTGGACATTGACAGGCTGGAATACCTCGGATATGACGGGAATCC 1200  
Db 1703 GCCCCCAACCATCTCTGGACATTGACAGGCTGGAATACCTCGGATATGACGGGAATCC 1762  
QY 1201 ATCTCTAAGCTGTCTGGACACGAGCGCGGTGAATCGGTTTCACTTGAAGAAGATG 1260  
Db 1763 ATCTCTAAGCTGTCTGGACACGAGCGCGGTGAATCGGTTTCACTTGAAGAAGATG 1822  
QY 1261 AGGCTCTGGCGGACTCCTTCTTGGTGAGAGGCAAGCTGTACACAGAGAGACAAT 1320  
Db 1823 AGGCTCTGGCGGACTCCTTCTTGGTGAGAGGCAAGCTGTACACAGAGAGACAAT 1882  
QY 1321 GACAAAGCTGACGCCACGAGGAGAACTTCTGCCMAATACAGAGCTGTGAAGGACCTG 1380  
Db 1883 GACAAAGCTGACGCCACGAGGAGAACTTCTGCCMAATACAGAGCTGTGAAGGACCTG 1942  
QY 1381 TGTCAAGCTGTGAGTACCAAGCGGCTGTGACAGCTGGGACAGAAAGTGGCAAGTGTG 1440  
Db 1943 TGTCAAGCTGTGAGTACCAAGCGGCTGTGACAGCTGGGACAGAAAGTGGCAAGTGTG 2002  
QY 1441 GAGAACCCACGGGGAAGCTGAAGCTGCATTAAGTGCAAGGGCCCCATCGGCTGGCGGC 1500  
Db 2003 GAGAACCCACGGGGAAGCTGAAGCTGCATTAAGTGCAAGGGCCCCATCGGCTGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGGGCAGGGCAGCGGCTGCAC 1560  
Db 2063 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGGGCAGGGCAGCGGCTGCAC 2122  
QY 1561 TGTGACAGCGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAAAAAACTTTCAAGAG 1620  
Db 2123 TGTGACAGCGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAAAAAACTTTCAAGAG 2182  
QY 1621 AAGTACAGGCCAGCTATGTCCGAGTGCCTCCATCGCTCAGTGGCCATCAGGTGGAC 1680  
Db 2183 AAGTACAGGCCAGCTATGTCCGAGTGCCTCCATCGCTCAGTGGCCATCAGGTGGAC 2242  
QY 1681 GGCAGGCTGTACCAAGCTAGGCTGGGTGTGCGGCCAGCCCGAAACCTCACCAAGCGG 1740  
Db 2243 GGCAGGCTGTACCAAGCTAGGCTGGGTGTGCGGCCAGCCCGAAACCTCACCAAGCGG 2302  
QY 1741 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 2362  
QY 1801 GGAGGCTTCCCGACTACTCAGCGGCCAACCCATTAAGTGGACATCGGTGTACATC 1860  
Db 2363 GGAGGCTTCCCGACTACTCAGCGGCCAACCCATTAAGTGGACATCGGTGTACATC 2422  
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTCAGGCTGGAA 1920  
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTCAGGCTGGAA 2482  
QY 1921 GACCAAGCTGCACATCGACACAGAGATTGAACCTCGAGAACAAAATTGAAGACCTG 1980  
Db 2483 GACCAAGCTGCACATCGACACAGAGATTGAACCTCGAGAACAAAATTGAAGACCTG 2542  
QY 1981 AGGGAAGTCCGAGTCACTGGAAGAAAAGCGGCCAGAGAAGTGTCACTGTGACAAAATC 2040  
Db 2543 AGGGAAGTCCGAGTCACTGGAAGAAAAGCGGCCAGAGAAGTGTCACTGTGACAAAATC 2602  
QY 2041 AGCTACCAACCCAGCAAAAGCGCGCTCAAGCAAGAGGCTCCAGTCTGCATCCTTTC 2100  
Db 2603 AGCTACCAACCCAGCAAAAGCGCGCTCAAGCAAGAGGCTCCAGTCTGCATCCTTTC 2662

QY 2101 AGGAAGGCGCTGCAAGAGAAAGGACAAAGTGTGGCTGTGTGGGGAGGAGAAAGCGCAAG 2160  
Db 2663 AGGAAGGCGCTGCAAGAGAAAGGACAAAGTGTGGCTGTGTGGGGAGGAGAAAGCGCAAG 2722  
QY 2161 AAACCTCGCAAGCTGCTCAAGCGCTGCAGAACAAACACACAGTGCAGCAGTGCAGGCGCTC 2220  
Db 2723 AAACCTCGCAAGCTGCTCAAGCGCTGCAGAACAAACACACAGTGCAGCAGTGCAGGCGCTC 2782  
QY 2221 ACCTGCTTCAACCCACGACAAACACAGCTGCGAGCGCGCTTCTCTGGACACATCGGCGCT 2280  
Db 2783 ACCTGCTTCAACCCACGACAAACACAGCTGCGAGCGCGCTTCTCTGGACACATCGGCGCT 2842  
QY 2281 TTCTGTGCTGTCACAGCGCCCAACAAATACAGTACTGTGTGCATGAGGACCATCAATGAG 2340  
Db 2843 TTCTGTGCTGTCACAGCGCCCAACAAATACAGTACTGTGTGCATGAGGACCATCAATGAG 2902  
QY 2341 ACTCACAATTTCTCTCTCTGTGAATTTGCAATTTGCAATTTCTAGAGTACTTTGATCTCAAC 2400  
Db 2903 ACTCACAATTTCTCTCTCTGTGAATTTGCAATTTCTAGAGTACTTTGATCTCAAC 2962  
QY 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGACAGGGGATGCTCAACAG 2460  
Db 2963 ACAGACCCCTACAGCTGATGATGAGTGAACACACACTGACAGGGGATGCTCAACAG 3022  
QY 2461 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 2520  
Db 3023 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 3082  
QY 2521 ACTCGAAACATGGAACCTGGGACTTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGTTT 2580  
Db 3083 ACTCGAAACATGGAACCTG-----GATGGAGGAGCTATGAGCAATAACAGGAGTTT 3133  
QY 2581 CAGGCTCGAAGTGGCCAGAAATGAGAGAGCTTCTTCAAAATCACTGGGACAACTGTGG 2640  
Db 3134 CAGGCTCGAAGTGGCCAGAAATGAGAGAGCTTCTTCAAAATCACTGGGACAACTGTGG 3193  
QY 2641 GAAGGCTGGGAAGTTTAAGAAACAAACAGAGGTGGACCTCCAAATAACATAGAGCATCACC 2700  
Db 3194 GAAGGCTGGGAAGTTTAAGAAACAAACAGAGGTGGACCTCCAAATAACATAGAGCATCACC 3253  
QY 2701 TGACTGCACAGGCMATGAATAAAACCATGTGGTGTATTTCCAGCAGACCTGTGTCTATTGGCC 2760  
Db 3254 TGACTGCACAGGCMATGAATAAAACCATGTGGTGTATTTCCAGCAGACCTGTGTCTATTGGCC 3313  
QY 2761 AGGAGGCTCAGAAAGCAACAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTCAGAAAGCAACAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAAATTTCAGGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATAAATTTCAGGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3433  
QY 2881 TACCTCACAGCTGCACAAATGCTTTTTCGTATCAAAAGTCCACCTACCTACCTCCC 2940  
Db 3434 TACCTCACAGCTGCACAAATGCTTTTTCGTATCAAAAGTCCACCTACCTACCTCCC 3493  
QY 2941 CCAGAGCTCACAAAGGAAACCGAGAGCGGAGAGAGATTTCTTTGGAAATTTCT 3000  
Db 3494 CCAGAGCTCACAAAGGAAACCGAGAGCGGAGAGAGATTTCTTTGGAAATTTCT 3553  
QY 3001 CCCAAGGCGGAAGTCAATTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3060  
Db 3554 CCCAAGGCGGAAGTCAATTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3613  
QY 3061 TCCTCTTATTCTTTTGTGTGTCAAAAGAGGAACCTTAAGAGCAGGACAGAGGCAACGT 3120  
Db 3614 TCCTCTTATTCTTTTGTGTGTGTCAAAAGAGGAACCTTAAGAGCAGGACAGAGGCAACGT 3673  
QY 3121 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGACAAATGAGTCACTAGCACAAAAGAGATGA 3180  
Db 3674 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGACAAATGAGTCACTAGCACAAAAGAGATGA 3733  
QY 3181 CATTTACTAGCACTATAAACCCCTGCTGCTGCTGAGAAACTGCTTCACTGTATATAT 3240

Db 3734 CATTACTAGCATATAAACCTGGTTGCTCTGAGAAACTGCCCTTCATTGTATATAT 3793  
QY 3241 GTGACTATTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAGAAATCCCAA 3300  
Db 3794 GTGACTATTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAGAAATCCCAA 3853  
QY 3301 TTTTCAGAGTGGTGGTCTCAATAAACGCTCTGTGGCCACTGTAAAGAAAAA 3353  
Db 3854 TTTTCAGAGTGGTGGTCTCAATAAACGCTCTGTGGCCACTGTAAAGAAAAA 3906

## RESULT 9

US-09-989-723-83

Sequence 83 Application US/09989723

Patent No. US20020072092A1

## GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnovers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

Acids Encoding the Same

FILE REFERENCE: P2730P1C62

CURRENT APPLICATION NUMBER: US/09/989,723

CURRENT FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-19

PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19



1643 AGGGCCCCAAGTGGAGCGGCTGTCTGATCCCAATCGTCTCAATTTGACCTG 1702  
1141 GCGCCACCAATCTTGACATTTGACGGCTTGACATATCTTGCGGATATGACGGAATTC 1200  
1703 GCGCCACCAATCTTGACATTTGACGGCTTGACATATCTTGCGGATATGACGGAATTC 1762  
1201 ATCTTCAAGCTCTGACACAGGAGCGCGGTGATCGTTTCACTTCAAAAAGAGATG 1260  
1763 ATCTTCAAGCTCTGACACAGGAGCGCGGTGATCGTTTCACTTCAAAAAGAGATG 1822  
1261 AGGGTCTGCGGGAATCTCTTTGTTGGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1320  
1823 AGGGTCTGCGGGAATCTCTTTGTTGGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1882  
1321 GACAGAGTGGACGCGCCAGGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAAGAGATG 1380  
1893 GACAGAGTGGACGCGCCAGGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAAGAGATG 1942  
1381 TGTACAGCTGTGAGTACCAAGACGGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440  
1943 TGTACAGCTGTGAGTACCAAGACGGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2002  
1441 GAGGACGCAACGGGAGAGCTGAGCTGATTAAGTGCAGAGGCGCCATGCGGCTGGCGGC 1500  
2003 GAGGACGCAACGGGAGAGCTGAGCTGATTAAGTGCAGAGGCGCCATGCGGCTGGCGGC 2062  
1501 AGCAGAGCCCTCTCAACCTCTGTGCCCAAGTACTACGGGAGGCGGACGAGGCGCTGCACC 1560  
2063 AGCAGAGCCCTCTCAACCTCTGTGCCCAAGTACTACGGGAGGCGGACGAGGCGCTGCACC 2122  
1561 TGTGACAGCGGGAATCAAGCTCAGCTGCGCGGACGCGGAGAAAACTCTTCAAGAAG 1620  
2123 TGTGACAGCGGGAATCAAGCTCAGCTGCGCGGACGCGGAGAAAACTCTTCAAGAAG 2182  
1621 AGTACAGGCGAGCTATGTGCGCAGTGTGCTCCATCGCTCAGTGGCCATCAGGTGCAC 1680  
2183 AGTACAGGCGAGCTATGTGCGCAGTGTGCTCCATCGCTCAGTGGCCATCAGGTGCAC 2242  
1681 GCGAGGCTGTACAGTGGGCTGTGATGCGCGCCAGCGCCGAAACCTCACCAAGCGG 1740  
2243 GCGAGGCTGTACAGTGGGCTGTGATGCGCGCCAGCGCCGAAACCTCACCAAGCGG 2302  
1741 CACTGCGCAGGCGCCCTTGAGACCAAGATGACAGAGATGTGGGACTTTCAGTGGCACT 1800  
2303 CACTGCGCAGGCGCCCTTGAGACCAAGATGACAGAGATGTGGGACTTTCAGTGGCACT 2362  
1801 GGAGGCTTTCGCACTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGCTATC 1860  
2363 GGAGGCTTTCGCACTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGCTATC 2422  
1861 CTAGAGACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCGCTGAAA 1920  
2423 CTAGAGACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCGCTGAAA 2482  
1921 GACCAAGCTGCACATCGACACAGATTTGAACCTTGCGAGACAAATTAAGAACCTG 1980  
2483 GACCAAGCTGCACATCGACACAGATTTGAACCTTGCGAGACAAATTAAGAACCTG 2542  
1981 AGGGAAGTCCGAGGTCACTCTGAAGAAAAGCGCGCAGAGAAATGTGACTGTACAAAATC 2040  
2543 AGGGAAGTCCGAGGTCACTCTGAAGAAAAGCGCGCAGAGAAATGTGACTGTACAAAATC 2602  
2041 AGTACACACACCGACACAAAGGCGCTCAAGCAGACAGAGCTCCAGTGTGCACTTTC 2100  
2603 AGTACACACACCGACACAAAGGCGCTCAAGCAGACAGAGCTCCAGTGTGCACTTTC 2662  
2101 AGGAAGGCGCTGCAAGAGAGAGAGAGTGTGGTGTGGGAGCAGAGCGCAGAGAG 2160  
2663 AGGAAGGCGCTGCAAGAGAGAGAGAGTGTGGTGTGGGAGCAGAGCGCAGAGAG 2722  
2161 AAACCTCCGAAAGCTGTCTCAAGCGCTTCAGAAACACGACAGTGCAGCATGCCAGGCGCTC 2220

2723 AAACCTCCGAAAGCTGTCTCAAGCGCTTCAGAGACACAGACACGTCGACATGCCAGGCGCTC 2782  
2221 ACCTGTCTTCACCACGACCAACAGACACTGGCAGACGCGCCCTTTCTGGACACTGGGCGCT 2280  
2783 ACCTGTCTTCACCACGACCAACAGACACTGGCAGACGCGCCCTTTCTGGACACTGGGCGCT 2842  
2281 TTCTGTGCTTGCACACGCGCCCAACATATACATGACTGTGTGATGAGGACCATCAATGAG 2340  
2843 TTCTGTGCTTGCACACGCGCCCAACATATACATGACTGTGTGATGAGGACCATCAATGAG 2902  
2341 ACTCACAATTTCTCTTCTGTGAAATTTGCAATGCGCTTCTTAGAGTACTTTGATCTCAAC 2400  
2903 ACTCACAATTTCTCTTCTGTGAAATTTGCAATGCGCTTCTTAGAGTACTTTGATCTCAAC 2962  
2401 ACAGACCCCTACAGCTGATGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2460  
2963 ACAGACCCCTACAGCTGATGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3022  
2461 CTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 2520  
3023 CTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 3082  
2521 ACTCGAAAATGGAACCTGGGACTTAAAGATGGAAGACTATGAGCAATACAGGCACTTT 2580  
3083 ACTCGAAAATGGAACCTGGGACTTAAAGATGGAAGACTATGAGCAATACAGGCACTTT 3133  
2581 CAGCGTGGAAAGTGGCCAGAAATGAGGAGACCTTTCTCCAAATCACTGGGACAACTGTGG 2640  
3134 CAGCGTGGAAAGTGGCCAGAAATGAGGAGACCTTTCTCCAAATCACTGGGACAACTGTGG 3193  
2641 GAAGGCTGGAAAGTGGAAACAAACAGAGGTGGACCTCCAAAACATAGAGGCACTACC 2700  
3194 GAAGGCTGGAAAGTGGAAACAAACAGAGGTGGACCTCCAAAACATAGAGGCACTACC 3253  
2701 TGACTGCACAGCAATGAAAACCAATGTTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 2760  
3254 TGACTGCACAGCAATGAAAACCAATGTTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 3313  
2761 AGGAGGCTTGAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAC 2820  
3314 AGGAGGCTTGAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAC 3373  
2821 CAGCAGGAGCAGAGATACTTCAGAAAGTCCATTTTGGCCCTGTCTTTGCTTTGGAATTA 2880  
3374 CAGCAGGAGCAGAGATACTTCAGAAAGTCCATTTTGGCCCTGTCTTTGCTTTGGAATTA 3433  
2881 TACCTCACAGCTGCACAAATGCAATTTTGGTATCAAAAAGTCCACCACTAACCTGCC 2940  
3434 TACCTCACAGCTGCACAAATGCAATTTTGGTATCAAAAAGTCCACCACTAACCTGCC 3493  
2941 CCAGAGGCTCACAAGGAAAACCGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3000  
3494 CCAGAGGCTCACAAGGAAAACCGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3553  
3001 CCAGAGGCGAAAAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTCTTAAA 3060  
3554 CCAGAGGCGAAAAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTCTTAAA 3613  
3061 TCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGAGCAGGCAACGT 3120  
3614 TCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGAGCAGGCAACGT 3673  
3121 GGAGGCGTGAACACAGTGCAGAGACGTTTGACAAATGATGATGATGATGATGATGATGAT 3180  
3674 GGAGGCGTGAACACAGTGCAGAGACGTTTGACAAATGATGATGATGATGATGATGATGAT 3733  
3181 CATTTACCTAGCACTATAAACCCTGTGCTTGAAGAACTGCTTCAATGATATAT 3240  
3734 CATTTACCTAGCACTATAAACCCTGTGCTTGAAGAACTGCTTCAATGATATAT 3793  
3241 GTGACTATTTTACATGATTAACATGGGAACTTTTGGGGGAACTTAATAAGAAATCCCAA 3300  
3794 GTGACTATTTTACATGATTAACATGGGAACTTTTGGGGGAACTTAATAAGAAATCCCAA 3853

Qy 3301 TTTTCAGGAGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3353  
Db 3854 TTTTCAGGAGTGGTGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3906

## RESULT 10

US-09-989-279-83  
/ Sequence 83, Application US/09989279  
/ Patent No. US20020072496A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Ashkenazi, Avi J.  
/ APPLICANT: Baker, Kevin P.  
/ APPLICANT: Botstein, David  
/ APPLICANT: Desnoyers, Luc  
/ APPLICANT: Baton, Dan L.  
/ APPLICANT: Ferrara, Napoleone  
/ APPLICANT: Fong, Sherman  
/ APPLICANT: Gerber, Hanspeter  
/ APPLICANT: Gerritsen, Mary E.  
/ APPLICANT: Goddard, Audrey  
/ APPLICANT: Godowski, Paul J.  
/ APPLICANT: Grimaldi, J. Christopher  
/ APPLICANT: Gurney, Austin L.  
/ APPLICANT: Kljavin, Ivar J.  
/ APPLICANT: Napier, Mary A.  
/ APPLICANT: Pan, James  
/ APPLICANT: Paoni, Nicholas F.  
/ APPLICANT: Roy, Margaret Ann  
/ APPLICANT: Stewart, Timothy A.  
/ APPLICANT: Tumas, Daniel  
/ APPLICANT: Watanabe, Colin K.  
/ APPLICANT: Williams, P. Mickey  
/ APPLICANT: Wood, William I.  
/ APPLICANT: Zhang, Zemin  
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
/ FILE REFERENCE: P2730P1C56  
/ CURRENT APPLICATION NUMBER: US/09/989,279  
/ CURRENT FILING DATE: 2001-11-19  
/ PRIOR APPLICATION NUMBER: 60/049787  
/ PRIOR FILING DATE: 1997-06-16  
/ PRIOR APPLICATION NUMBER: 60/062250  
/ PRIOR FILING DATE: 1997-10-17  
/ PRIOR APPLICATION NUMBER: 60/065186  
/ PRIOR FILING DATE: 1997-11-12  
/ PRIOR APPLICATION NUMBER: 60/065311  
/ PRIOR FILING DATE: 1997-11-13  
/ PRIOR APPLICATION NUMBER: 60/066770  
/ PRIOR FILING DATE: 1997-11-24  
/ PRIOR APPLICATION NUMBER: 60/075945  
/ PRIOR FILING DATE: 1998-02-25  
/ PRIOR APPLICATION NUMBER: 60/078910  
/ PRIOR FILING DATE: 1998-03-20  
/ PRIOR APPLICATION NUMBER: 60/083322  
/ PRIOR FILING DATE: 1998-04-28  
/ PRIOR APPLICATION NUMBER: 60/084600  
/ PRIOR FILING DATE: 1998-05-07  
/ PRIOR APPLICATION NUMBER: 60/087106  
/ PRIOR FILING DATE: 1998-05-28  
/ PRIOR APPLICATION NUMBER: 60/087607  
/ PRIOR FILING DATE: 1998-06-02  
/ PRIOR APPLICATION NUMBER: 60/087609  
/ PRIOR FILING DATE: 1998-06-02  
/ PRIOR APPLICATION NUMBER: 60/087759  
/ PRIOR FILING DATE: 1998-06-02  
/ PRIOR APPLICATION NUMBER: 60/087827  
/ PRIOR FILING DATE: 1998-06-03  
/ PRIOR APPLICATION NUMBER: 60/088021  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088025  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088026

/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088028  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088029  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088030  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088033  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088326  
/ PRIOR FILING DATE: 1998-06-04  
/ PRIOR APPLICATION NUMBER: 60/088167  
/ PRIOR FILING DATE: 1998-06-05  
/ PRIOR APPLICATION NUMBER: 60/088202  
/ PRIOR FILING DATE: 1998-06-05  
/ PRIOR APPLICATION NUMBER: 60/088212  
/ PRIOR FILING DATE: 1998-06-05  
/ PRIOR APPLICATION NUMBER: 60/088217  
/ PRIOR FILING DATE: 1998-06-05  
/ PRIOR APPLICATION NUMBER: 60/088655  
/ PRIOR FILING DATE: 1998-06-09  
/ PRIOR APPLICATION NUMBER: 60/088734  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088738  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088742  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088810  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088824  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088826  
/ PRIOR FILING DATE: 1998-06-10  
/ PRIOR APPLICATION NUMBER: 60/088858  
/ PRIOR FILING DATE: 1998-06-11  
/ PRIOR APPLICATION NUMBER: 60/088861  
/ PRIOR FILING DATE: 1998-06-11  
/ PRIOR APPLICATION NUMBER: 60/088876  
/ PRIOR FILING DATE: 1998-06-11  
/ PRIOR APPLICATION NUMBER: 60/089105  
/ PRIOR FILING DATE: 1998-06-12  
/ PRIOR APPLICATION NUMBER: 60/089440  
/ PRIOR FILING DATE: 1998-06-16  
/ PRIOR APPLICATION NUMBER: 60/089512  
/ PRIOR FILING DATE: 1998-06-16  
/ PRIOR APPLICATION NUMBER: 60/089514  
/ PRIOR FILING DATE: 1998-06-16  
/ PRIOR APPLICATION NUMBER: 60/089532  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089538  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089598  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089599  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089600  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089653  
/ PRIOR FILING DATE: 1998-06-17  
/ PRIOR APPLICATION NUMBER: 60/089801  
/ PRIOR FILING DATE: 1998-06-18  
/ PRIOR APPLICATION NUMBER: 60/089907  
/ PRIOR FILING DATE: 1998-06-18  
/ PRIOR APPLICATION NUMBER: 60/089908  
/ PRIOR FILING DATE: 1998-06-18  
/ PRIOR APPLICATION NUMBER: 60/089947  
/ PRIOR FILING DATE: 1998-06-19  
/ PRIOR APPLICATION NUMBER: 60/089948  
/ PRIOR FILING DATE: 1998-06-19  
/ PRIOR APPLICATION NUMBER: 60/089952  
/ PRIOR FILING DATE: 1998-06-19  
/ PRIOR APPLICATION NUMBER: 60/090246  
/ PRIOR FILING DATE: 1998-06-22







RESULT 11  
US-09-989-727-83  
Sequence 83, Application US/09899727  
Patent No. US20020072497A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
Acids Encoding the Same  
FILE REFERENCE: P2730P1C65  
CURRENT APPLICATION NUMBER: US/09/989,727  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23

/	Prior Application Number: 60/090355																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			</
---	-------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Db 1823 AGGCTCTGGGGGACCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1882  
Qy 1321 GACAAGGTGAGACGCGCCAGGAGGAGAACTTTCTGCCCAAGTACCAGAGTGTTGAGAGGACCTG 1380  
Db 1883 GACAAGGTGAGACGCGCCAGGAGGAGAACTTTCTGCCCAAGTACCAGAGTGTTGAGAGGACCTG 1942  
Qy 1381 TGTGAGCGTGTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAGTGGCAGTGTGTG 1440  
Db 1943 TGTGAGCGTGTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAGTGGCAGTGTGTG 2002  
Qy 1441 GAGGACGCCACGGGGAAGCTGAAGCTGCATAGTGAAGGCGCCCACTGGCGCTGGCGGCG 1500  
Db 2003 GAGGACGCCACGGGGAAGCTGAAGCTGCATAGTGAAGGCGCCCACTGGCGCTGGCGGCG 2062  
Qy 1501 AGCAGAGCCCTCTCCAACTCGTGCCTGTCCTCCCAAGTACTACGGGAGGCGGACGAGGCTGCACC 1560  
Db 2063 AGCAGAGCCCTCTCCAACTCGTGCCTGTCCTCCCAAGTACTACGGGAGGCGGACGAGGCTGCACC 2122  
Qy 1561 TGTGACAGCGGGACTACAGCTCAGCTGGCGGAGCGCGGAAACCTCTTCAAGAG 1620  
Db 2123 TGTGACAGCGGGACTACAGCTCAGCTGGCGGAGCGCGGAAACCTCTTCAAGAG 2182  
Qy 1621 AAGTACAGGCGACGCTATGTCGCGAGTGCCTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680  
Db 2183 AAGTACAGGCGACGCTATGTCGCGAGTGCCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242  
Qy 1681 GGCAGGCTGTACCAAGTGGCTGGTGTGTCGCGCCGAGCGCCGAAACCTCTCCAAAGCGG 1740  
Db 2243 GGCAGGCTGTACCAAGTGGCTGGTGTGTCGCGCCGAGCGCCGAAACCTCTCCAAAGCGG 2302  
Qy 1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACT 2362  
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTAAGTGCACATCGTGTCTACATC 1860  
Db 2363 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTAAGTGCACATCGTGTCTACATC 2422  
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACTGGACCTGTACAGTCCCTGCGGCTGGAAA 1920  
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACTGGACCTGTACAGTCCCTGCGGCTGGAAA 2482  
Qy 1921 GACCACAAGTGTGCATCGACACAGATTTGAAACCTCTGACAGCAAAATTAAGAACCTG 1980  
Db 2483 GACCACAAGTGTGCATCGACACAGATTTGAAACCTCTGACAGCAAAATTAAGAACCTG 2542  
Qy 1981 AGGGAAGTCCGAGTCACTGAGAAAGGCGCCAGAGAAATGTGACTGTCACAAAATC 2040  
Db 2543 AGGGAAGTCCGAGTCACTGAGAAAGGCGCCAGAGAAATGTGACTGTCACAAAATC 2602  
Qy 2041 AGCTACACACCCAGCAAGGCGGCTCAAGCAGAGGCTCCAGTCTGCACTCCCTTC 2100  
Db 2603 AGCTACACACCCAGCAAGGCGGCTCAAGCAGAGGCTCCAGTCTGCACTCCCTTC 2662  
Qy 2101 AGGAAGGCTTGCAGAGAGGCAAGGCTGTGTTGGGGAGCAGAGCGCAAGAG 2160  
Db 2663 AGGAAGGCTTGCAGAGAGGCAAGGCTGTGTTGGGGAGCAGAGCGCAAGAG 2722  
Qy 2161 AAATCCGCAAGCTGTCAAGCGCTGCAGAACACGACAGTGCAGATGCCAGGCTC 2220  
Db 2723 AAATCCGCAAGCTGTCAAGCGCTGCAGAACACGACAGTGCAGATGCCAGGCTC 2782  
Qy 2221 ACGTGCTTCAACCCAGCAACCCAGCTGGCAGCGGCTTCTTGGACACTGGGGCT 2280  
Db 2783 ACGTGCTTCAACCCAGCAACCCAGCTGGCAGCGGCTTCTTGGACACTGGGGCT 2842  
Qy 2281 TTCTGTGCTGTGACAGGCGCAACATPAACAGTATCTGGTGCAATGAGGACCATCAATGAG 2340  
Db 2843 TTCTGTGCTGTGACAGGCGCAACATPAACAGTATCTGGTGCAATGAGGACCATCAATGAG 2902  
Qy 2341 ACTCACAATTTCTTCTGTGATTTGCACTGGCTTCTTGTAGACTTTGATCTCAAC 2400

Db 2903 ACTCACAATTTCTTCTGTGAATTTGCACCTGGCTTCTTAGAGTACTTTGATCTCAAC 2962  
Qy 2401 ACAGACCCCTACCAAGCTGATCAATGCACTGAGCAAGGATGTCTCAACCCAG 2460  
Db 2963 ACAGACCCCTACCAAGCTGATCAATGCACTGAGCAAGGATGTCTCAACCCAG 3022  
Qy 2461 CTACAGCTACAGCTCATGAGAGCTGAGAGCTGCAAGGGTTACAAGCGAGTGTAAACCCCGG 2520  
Db 3023 CTACAGCTACAGCTCATGAGAGCTGAGAGCTGCAAGGGTTACAAGCGAGTGTAAACCCCGG 3082  
Qy 2521 ACTCGAATACATGAGAGCTTGAAGATGAGAGAGCTATGAGCAATACAGGACGTTT 2580  
Db 3083 ACTCGAATACATGAGAGCTTGAAGATGAGAGAGCTTGAAGCAATACAGGACGTTT 3133  
Qy 2581 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGCAACTGTGG 2640  
Db 3134 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGCAACTGTGG 3193  
Qy 2641 GAAGCTGGGAGGTTAAGAAACAACAGAGGTGACCTTCCAAACATAGAGGCACTCACC 2700  
Db 3194 GAAGCTGGGAGGTTAAGAAACAACAGAGGTGACCTTCCAAACATAGAGGCACTCACC 3253  
Qy 2701 TGACTGCAAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760  
Db 3254 TGACTGCAAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313  
Qy 2761 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 2820  
Db 3314 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 3373  
Qy 2821 CAGCAGGAGCAGAGATTAATTTGAGAGTCCATTTTGGCCCTGCTTTGGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATTAATTTGAGAGTCCATTTTGGCCCTGCTTTGGGATTA 3433  
Qy 2881 TACTCTACAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACCTAACCCCTCC 2940  
Db 3434 TACTCTACAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACCTAACCCCTCC 3493  
Qy 2941 CCAGAGCTCAAAAGGAAACCGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000  
Db 3494 CCAGAGCTCAAAAGGAAACCGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553  
Qy 3001 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTGTTCTAAA 3060  
Db 3554 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTGTTCTAAA 3613  
Qy 3061 TCCTCTTATTTCTTTGGTTGTCTCAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3120  
Db 3614 TCCTCTTATTTCTTTGGTTGTCTCAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3673  
Qy 3121 GGAGAGGCTGAAACAGTGCAGAGACGTTTGACAAATGAGTCAGTACGACAAAGAGATGA 3180  
Db 3674 GGAGAGGCTGAAACAGTGCAGAGACGTTTGACAAATGAGTCAGTACGACAAAGAGATGA 3733  
Qy 3181 CATTTACCTAGCATATAAACCTTGGTTGCTCTTGAAGAACTGCGCTTCAATTTGATATAT 3240  
Db 3734 CATTTACCTAGCATATAAACCTTGGTTGCTCTTGAAGAACTGCGCTTCAATTTGATATAT 3793  
Qy 3241 GTGACTATTTACATGTAATCAACATGGGACTTTTAGGGGAACTTAATAGAAATCCCAA 3300  
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACTTAATAGAAATCCCAA 3853  
Qy 3301 TTTTCAGGAGTGGTGTCTCAATTAAGCTCTGTGGCCAGTGTAAAAAGAAAA 3353  
Db 3854 TTTTCAGGAGTGGTGTCTCAATTAAGCTCTGTGGCCAGTGTAAAAAGAAAA 3906

## RESULT 12

US-09-989-731-83  
; Sequence 83, Application US/09989731  
; Patent No. US2020103125A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin I.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC70  
CURRENT APPLICATION NUMBER: US/09/989,731  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24

[illegible]



1381 TGTCAGCGTCTGAGTACAGACGGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440  
1943 TGTCAGCGTCTGAGTACAGACGGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2002  
1441 GAGGACGCCACGGGGAAGCTGAAGCTGCAATAAGTGAAGGCGCCCATCGGCTGGCGGC 1500  
2003 GAGGACGCCACGGGGAAGCTGAAGCTGCAATAAGTGAAGGCGCCCATCGGCTGGCGGC 2062  
1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGCGCAGGCGAGGCGCTGCACC 1560  
2063 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGCGCAGGCGAGGCGCTGCACC 2122  
1561 TGTGACAGCGGGAAGTACAAGCTGAGCTGCGCGACGCGCGGAAAGAACTCTTCAAGAG 1620  
2123 TGTGACAGCGGGAAGTACAAGCTGAGCTGCGCGACGCGCGGAAAGAACTCTTCAAGAG 2182  
1621 AGTACAGCGCAGCTATGTCGCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 1680  
2183 AGTACAGCGCAGCTATGTCGCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 2242  
1681 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGCGCCAGCCCGGAAACCTCACCAAGCGG 1740  
2243 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGCGCCAGCCCGGAAACCTCACCAAGCGG 2302  
1741 CACTGCGCAGGCGCCCTGAGGACCGAAGATGACAAAGATGCTGGGGAATGAGTGGCACT 1800  
2303 CACTGCGCAGGCGCCCTGAGGACCGAAGATGACAAAGATGCTGGGGAATGAGTGGCACT 2362  
1801 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGCGCCAGCCCGGAAACCTCACCAAGCGG 1860  
2363 GGCAGGGTGTACCACTAGGCGCTGGGTGATGCGCGCCAGCCCGGAAACCTCACCAAGCGG 2422  
1861 CTAGAGAACGACAGCTCCAGTGTGACCTGGACCTGTACAACTCCCTGCGAGCCCTGGAAA 1920  
2423 CTAGAGAACGACAGCTCCAGTGTGACCTGGACCTGTACAACTCCCTGCGAGCCCTGGAAA 2482  
1921 GACCAAGCTGCACATCGACACGAGATTGAAACCTCGCAGAACAAATTAAGAACCTG 1980  
2483 GACCAAGCTGCACATCGACACGAGATTGAAACCTCGCAGAACAAATTAAGAACCTG 2542  
1981 AGGGAAGTCCGAGGTCACTGGAAGAAAAGCGGCGCAGAGAAATGTGACTGTACAAAATC 2040  
2543 AGGGAAGTCCGAGGTCACTGGAAGAAAAGCGGCGCAGAGAAATGTGACTGTACAAAATC 2602  
2041 AGCTACACACCCAGCACAAAGCGGCTCAGACAGAGGCTCAGTCTGATCCTTTC 2100  
2603 AGCTACACACCCAGCACAAAGCGGCTCAGACAGAGGCTCAGTCTGATCCTTTC 2662  
2101 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTGCGGAGCAGAGCGCAAGAG 2160  
2663 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTGCGGAGCAGAGCGCAAGAG 2722  
2161 AAACCTCGGAAGTGTCTAAGCGCTTCAGAACAAAGCAGATGCGAGATGCGAGGCTC 2220  
2723 AAACCTCGGAAGTGTCTAAGCGCTTCAGAACAAAGCAGATGCGAGATGCGAGGCTC 2782  
2221 AGCTGCTTCAACCCAGACACAGCACTGCGAGAGCGGCGCTTCTGAGCACTGGGCGCT 2280  
2783 AGCTGCTTCAACCCAGACACAGCACTGCGAGAGCGGCGCTTCTGAGCACTGGGCGCT 2842  
2281 TTCTGTGCTGCAACGCGCCCAACATAACAGTACTGCTGATGAGGACCATCAATGAG 2340  
2843 TTCTGTGCTGCAACGCGCCCAACATAACAGTACTGCTGATGAGGACCATCAATGAG 2902  
2341 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGCTTCTAGAGTACTTTGATCTCAAC 2400  
2903 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGCTTCTAGAGTACTTTGATCTCAAC 2962  
2401 ACAGACCCCTTACAGCTGATGATGAGTGAAGACACACTGGACAGGGAATGCTCTCAAC 2460  
2963 ACAGACCCCTTACAGCTGATGATGAGTGAAGACACACTGGACAGGGAATGCTCTCAAC 3022

2461 CTACAGCTCAGCTCATGAGCTGAGAGCTGCAAGCGGTTTCAAGCAGTGTAAACCCCGG 2520  
3023 CTACAGCTCAGCTCATGAGCTGAGAGCTGCAAGCGGTTTCAAGCAGTGTAAACCCCGG 3082  
2521 ACTCGAAAACATGACCTTGGGACTTAAAGATGAGGAGCTATGAGCAATACAGGCAAGTTT 2580  
3083 ACTCGAAAACATGACCTG-----GATGGAGGAAGCTATGAGCAATACAGGCAAGTTT 3133  
2581 CAGCGTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACCTGTGG 2640  
3134 CAGCGTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACCTGTGG 3193  
2641 GAAGGCTGGGAAGTTAAGAAAACACAGAGGTGGACCTCCAAAACATAGAGGATCAAC 2700  
3194 GAAGGCTGGGAAGTTAAGAAAACACAGAGGTGGACCTCCAAAACATAGAGGATCAAC 3253  
2701 TGACTGCACAGGCAATGAAGAAACCATGTGGGTGATTTCCAGAGACCTGTCTATTGGCC 2760  
3254 TGACTGCACAGGCAATGAAGAAACCATGTGGGTGATTTCCAGAGACCTGTCTATTGGCC 3313  
2761 AGGAGGCTGAGAAAACGACGACCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
3314 AGGAGGCTGAGAAAACGACGACCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
2821 CAGCAGGACGAGATTAACCTTCAAGGAAGTCCATTTTGGCCCTGTCTTTGCTTTGATTA 2880  
3374 CAGCAGGACGAGATTAACCTTCAAGGAAGTCCATTTTGGCCCTGTCTTTGCTTTGATTA 3433  
2881 TACTCACCAGCTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCTCC 2940  
3434 TACTCACCAGCTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCTCC 3493  
2941 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000  
3494 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3553  
3001 CCCAAGCGGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAA 3060  
3554 CCCAAGCGGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAA 3613  
3061 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAACTAAGAAAGCAGGAGCAAGCT 3120  
3614 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAACTAAGAAAGCAGGCAAGCT 3673  
3121 GGAGGCTGAAAACAGTGCAGAGAGCTTGAATGAGTCAAGTCAAGAGAGATGA 3180  
3674 GGAGGCTGAAAACAGTGCAGAGAGCTTGAATGAGTCAAGTCAAGAGAGATGA 3733  
3181 CATTTACCTAGCACTATAAACCTTGTCTCTGAAGAACTGCTTCAATGTATATAT 3240  
3734 CATTTACCTAGCACTATAAACCTTGTCTCTGAAGAACTGCTTCAATGTATATAT 3793  
3241 GTGACTATTTACATGTAATCAATGAGGAACTTTTAGGGAACTTAATAGAAATCCCAA 3300  
3794 GTGACTATTTACATGTAATCAATGAGGAACTTTTAGGGAACTTAATAGAAATCCCAA 3853  
3301 TTTTCAGGAGTGGTGTGCAATTAACAGCTCTGTGGCCAGTGTAAAAGAAAAA 3353  
3854 TTTTCAGGAGTGGTGTGCAATTAACAGCTCTGTGGCCAGTGTAAAAGAAAAA 3906

## RESULT 13

US-09-989-732-83  
; Sequence 83, Application US/09989732  
; Patent No. US20020123463A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pat, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C57  
CURRENT APPLICATION NUMBER: US/09/989,732  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-15  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24

[illegible]

Db 2003 GAGACGCCACCGGGAAGCTGAAGCTGCATAAGTGAAGGCCGCCCTGCGCTGGGCGGC 2062  
Qy 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 1560  
Db 2063 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 2122  
Qy 1561 TGTGACAGCGGGGACTACAAGCTCAGCTCGGCGGACGCGCGGAAAAAATCTTCAAGAAG 1620  
Db 2123 TGTGACAGCGGGGACTACAAGCTCAGCTCGGCGGACGCGCGGAAAAAATCTTCAAGAAG 2182  
Qy 1621 AAGTACAAGGCGACGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680  
Db 2183 AAGTACAAGGCGACGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242  
Qy 1681 GGCAGGGTGTACCACTGAGGCTGGGTGATGCGCGCCCGAGCCCGGAAACCTTCAACAAGCGG 1740  
Db 2243 GGCAGGGTGTACCACTGAGGCTGGGTGATGCGCGCCCGAGCCCGGAAACCTTCAACAAGCGG 2302  
Qy 1741 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAAAGATGGTGGGACTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAAAGATGGTGGGACTTCAGTGGCACT 2362  
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCGCCCAACCCCAATTAAGTGCACATCGGTCTACATC 1860  
Db 2363 GGAGGCTTCCCGACTACTCAGCGCGCCCAACCCCAATTAAGTGCACATCGGTCTACATC 2422  
Qy 1861 CTAGAGAACACACAGTCCAGTGTGACTGGACCTGTGTAAGTCCCTGCGAGGCTCGAAA 1920  
Db 2423 CTAGAGAACACACAGTCCAGTGTGACTGGACCTGTGTAAGTCCCTGCGAGGCTCGAAA 2482  
Qy 1921 GACCACAAGCTGCACATCGACACGAGATGAAACCTCGCAGAACCAAAATTAAGAACCTG 1980  
Db 2483 GACCACAAGCTGCACATCGACACGAGATGAAACCTCGCAGAACCAAAATTAAGAACCTG 2542  
Qy 1981 AGGAAGTCCGAGTGCACCTGAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2040  
Db 2543 AGGAAGTCCGAGTGCACCTGAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2602  
Qy 2041 AGCTACACACCCAGCAGCAAGGCGCCCTCAAGCAGAGGCTCAGTCTGCATCCTTTC 2100  
Db 2603 AGCTACACACCCAGCAGCAAGGCGCCCTCAAGCAGAGGCTCAGTCTGCATCCTTTC 2662  
Qy 2101 AGGAAGGCTCCGAGAGAGGCAAGGTGTGCTGTGCGGAGCAGAGCGCAAGAG 2160  
Db 2663 AGGAAGGCTCCGAGAGAGGCAAGGTGTGCTGTGCGGAGCAGAGCGCAAGAG 2722  
Qy 2161 AAATCCGAGGCTGCTCAAGCGGCTGCGAACAACGACAGTGCAGATGCGAGGCTC 2220  
Db 2723 AAATCCGAGGCTGCTCAAGCGGCTGCGAACAACGACAGTGCAGATGCGAGGCTC 2782  
Qy 2221 ACGTGCTTCCACCAAGCAACCGACTGCGACAGCGGCTTTCGAGACATCGGGCT 2280  
Db 2783 ACGTGCTTCCACCAAGCAACCGACTGCGACAGCGGCTTTCGAGACATCGGGCT 2842  
Qy 2281 TTCTGTGCTGCAACGAGCGCAACAATAACGCTACTGTGCTGATGAGGACCATCAATGAG 2340  
Db 2843 TTCTGTGCTGCAACGAGCGCAACAATAACGCTACTGTGCTGATGAGGACCATCAATGAG 2902  
Qy 2341 ACTCAAAATTCCTCTCTGTGATTTGCAATGCACTGGCTTCTAGAGTACTTGTATCTCAAC 2400  
Db 2903 ACTCAAAATTCCTCTCTGTGATTTGCAATGCACTGGCTTCTAGAGTACTTGTATCTCAAC 2962  
Qy 2401 ACAGACCCCTACGAGTGTGATGAGTGCAGTGAAACACACTGACAGGATGTCTCAACAG 2460  
Db 2963 ACAGACCCCTACGAGTGTGATGAGTGCAGTGAAACACACTGACAGGATGTCTCAACAG 3022  
Qy 2461 CTACAGTACAGTCTGAGGCTGAGAGCTGCAAGGTTAAGAGTGTAAACCCCGG 2520  
Db 3023 CTACAGTACAGTCTGAGGCTGAGAGCTGCAAGGTTAAGAGTGTAAACCCCGG 3082  
Qy 2521 ACTCGAAACATGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCACTTT 2580

Db 3083 ACTCGAAACATGGACCTG-----GATGGAGGAAGCTATGAGCAATACAGGCACTTT 3133  
Qy 2581 CAGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACTGTGG 2640  
Db 3134 CAGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACTGTGG 3193  
Qy 2641 GAAGCTCGGAAGGTTAAGAAACCAAGAGGTGACCTTCCAAAAACATAGAGGCACTCAC 2700  
Db 3194 GAAGCTCGGAAGGTTAAGAAACCAAGAGGTGACCTTCCAAAAACATAGAGGCACTCAC 3253  
Qy 2701 TCATGTCACAGGCAATGAAGAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 2760  
Db 3254 TCATGTCACAGGCAATGAAGAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 3313  
Qy 2761 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
Qy 2821 CAGCAGGAGCAGAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 3433  
Qy 2881 TACCTCACACCTGTCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTTCCC 2940  
Db 3434 TACCTCACACCTGTCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTTCCC 3493  
Qy 2941 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000  
Db 3494 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553  
Qy 3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGAAAGCAGTCTGTCTTAAA 3060  
Db 3554 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGAAAGCAGTCTGTCTTAAA 3613  
Qy 3061 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3120  
Db 3614 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3673  
Qy 3121 GGAGAGCTGAAACAGTGCAGAGACGTTTGAACATGAGTCACTAGCAGCAAAAGAGATGA 3180  
Db 3674 GGAGAGCTGAAACAGTGCAGAGACGTTTGAACATGAGTCACTAGCAGCAAAAGAGATGA 3733  
Qy 3181 CATTTACTAGCATATAAAACCTTGTGCTCTGAAAGAACTGCTTTCATTGTATATAT 3240  
Db 3734 CATTTACTAGCATATAAAACCTTGTGCTCTGAAAGAACTGCTTTCATTGTATATAT 3793  
Qy 3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTCTAATAAGAAATCCCAA 3300  
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTCTAATAAGAAATCCCAA 3853  
Qy 3301 TTTTCAGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAAGAAAAA 3353  
Db 3854 TTTTCAGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAAGAAAAA 3906

RESULT 14  
US-09-991-073-83  
; Sequence 83, Application US/09991073  
; Patent No. US20202127576A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Daniel  
APPLICANT: Tamas, Colin K.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C15  
CURRENT APPLICATION NUMBER: US/09/991,073  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24

;	Prior Application Number: 60/090542																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
---	-------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



1561 TGTGACAGCGGGGACTACAAGCTCAGCCTGCGCGGAGCGCCGGAATAAATCTTCAAGAAG 1620  
1562 TGTGACAGCGGGGACTACAAGCTCAGCCTGCGCGGAGCGCCGGAATAAATCTTCAAGAAG 2182  
1621 AAGTACAAGCGGAGCTATGTCGCGAGTGCCTCCATTCGCTCAGTGGGAGTGCAGTGGGAG 1680  
2183 AAGTACAAGCGGAGCTATGTCGCGAGTGCCTCCATTCGCTCAGTGGGAGTGCAGTGGGAG 2242  
1681 GGCAGGGGTACACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1740  
2243 GGCAGGGGTACACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 2302  
1741 CACTGGCAGGGGCCCCCTGAGGACCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 1800  
2303 CACTGGCAGGGGCCCCCTGAGGACCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 2362  
1801 GGAGGCGCTTCCGAGTACTCAGCGCCCAACCCCTTAAAGTGCAGATCGGTGCTACATC 1860  
2363 GGAGGCGCTTCCGAGTACTCAGCGCCCAACCCCTTAAAGTGCAGATCGGTGCTACATC 2422  
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGCGAGGCTGGAAA 1920  
2423 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGCGAGGCTGGAAA 2482  
1921 GACCAAGCTGCATCGACACAGATTTGAACCTTGCAGAACCAAAATTAAGAACCTG 1980  
2483 GACCAAGCTGCATCGACACAGATTTGAACCTTGCAGAACCAAAATTAAGAACCTG 2542  
1981 AGGGAAGTCCGAGGTCACCTGAAGAAAAGCGGCGCAGAGGAATGTGACTGTCACAAAATC 2040  
2543 AGGGAAGTCCGAGGTCACCTGAAGAAAAGCGGCGCAGAGGAATGTGACTGTCACAAAATC 2602  
2041 AGCTACACACCCAGGACAAAGGCGGCTCAAGACAGAGGCTCGAGTGTGATCTCTTTC 2100  
2603 AGCTACACACCCAGGACAAAGGCGGCTCAAGACAGAGGCTCGAGTGTGATCTCTTTC 2662  
2101 AGCAAGGGGCTCAAGAGAGGACAAAGTGTGGCTGTGGCGGAGCAGAGCGCAAGAG 2160  
2663 AGCAAGGGGCTCAAGAGAGGACAAAGTGTGGCTGTGGCGGAGCAGAGCGCAAGAG 2722  
2161 AAATCCGCAAGTGTCAAGCGCTGACAGAACCAAGACAGTGCAGATGCCAGGCTC 2220  
2723 AAATCCGCAAGTGTCAAGCGCTGACAGAACCAAGACAGTGCAGATGCCAGGCTC 2782  
2221 AGTGTCTTCAACCCAGACACAGCAGTGGCAGCGGCGCTTCTGACACAGTGGGCT 2280  
2783 AGTGTCTTCAACCCAGACACAGCAGTGGCAGCGGCGCTTCTGACACAGTGGGCT 2842  
2281 TTCTGTGCTGCACAGCGCCCAACCAATACAGTACTGTGATGAGGACCAATCAATGAG 2340  
2843 TTCTGTGCTGCACAGCGCCCAACCAATACAGTACTGTGATGAGGACCAATCAATGAG 2902  
2341 ACTCAAAATTTCTTCTGTGAAATTTGCAATGCTGCTTCTAGAGTACTTTGATCTCAAC 2400  
2903 ACTCAAAATTTCTTCTGTGAAATTTGCAATGCTGCTTCTAGAGTACTTTGATCTCAAC 2962  
2401 ACAGACCCCTTACAGCTGATGAATGAGTGAACACACACTGGAAGGAGTGTCTCAACAG 2460  
2963 ACAGACCCCTTACAGCTGATGAATGAGTGAACACACACTGGAAGGAGTGTCTCAACAG 3022  
2461 CTACAGTACAGCTATGAGCTGAGGAGCTGCAAGGGTTAACAAGTGTAAACCCCGG 2520  
3023 CTACAGTACAGCTATGAGCTGAGGAGCTGCAAGGGTTAACAAGTGTAAACCCCGG 3082  
2521 ACTCGAAACATGAGCTGGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 2580  
3083 ACTCGAAACATGAGCTGGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 3133  
2581 CAGCTGCAAGAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640  
3134 CAGCTGCAAGAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193

QY 2641 GAAGGCTGGGAAGTTAAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGATCACC 2700  
Db 3194 GAAGGCTGGGAAGTTAAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGATCACC 3253  
QY 2701 TGACTGCAAGGCAATGAATAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760  
Db 3254 TGACTGCAAGGCAATGAATAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313  
QY 2761 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAAATCTCAGGAAGTCAATTTTCCCTGCTGCTTTGCTTGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATAAATCTCAGGAAGTCAATTTTCCCTGCTGCTTTGCTTGGATTA 3433  
QY 2881 TACCTCAGCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTGCC 2940  
Db 3434 TACCTCAGCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTGCC 3493  
QY 2941 CCAAGAGCTCAAAAAGGAAAACGAGAGAGCGAGAGATTTCTTGGAAATTTCT 3000  
Db 3494 CCAAGAGCTCAAAAAGGAAAACGAGAGAGCGAGAGATTTCTTGGAAATTTCT 3553  
QY 3001 CCCAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGAAAAGCAGTCTGTTCTAAA 3060  
Db 3554 CCCAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGAAAAGCAGTCTGTTCTAAA 3613  
QY 3061 TCTCTTATTTCTTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGGAGCAGAGCAACGT 3120  
Db 3614 TCTCTTATTTCTTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGGAGCAGAGCAACGT 3673  
QY 3121 GGAGGCTGAAAACAGTGCAGAGAGCTTTGACATGATCAGTAGCACAAGAGAGATGA 3180  
Db 3674 GGAGGCTGAAAACAGTGCAGAGAGCTTTGACATGATCAGTAGCACAAGAGAGATGA 3733  
QY 3181 CATTTACCTAGCAGTATTAACCTTGGTCTGCTCTCAAGAACTGCTTCTTGTATATAT 3240  
Db 3734 CATTTACCTAGCAGTATTAACCTTGGTCTGCTCTCAAGAACTGCTTCTTGTATATAT 3793  
QY 3241 GTGACTATTTAGATGATTAATCAACATGGGAACTTTTAGGGAACTTAAGAAATCCCAA 3300  
Db 3794 GTGACTATTTAGATGATTAATCAACATGGGAACTTTTAGGGAACTTAAGAAATCCCAA 3853  
QY 3301 TTTTACGAGTGTGTGTCAATAAAGCTCTGTGCGCAGTGTAAAGAAAAA 3353  
Db 3854 TTTTACGAGTGTGTGTCAATAAAGCTCTGTGCGCAGTGTAAAGAAAAA 3906

## RESULT 15

US-09-990-442-83  
; Sequence 83, Application US/09990442  
; Patent No. US20020132252A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC8  
CURRENT APPLICATION NUMBER: US/09/990,442  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25

		Query Match		98.8%; Score 3323.4; DB 10; Length 3906;
		Best Local Similarity		99.7%; Pred. No. 0;
		Matches 3343; Conservative		0; Indels 9; Gaps 1;
QY	Db	1	GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAAGAGATGGCCCGCCGAGC	60
		563	GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAAGAGATGGCCCGCCGAGC	622
QY	Db	61	CTCGTCTGTGCTGTGTCGCAACTGTTCTCCCTGCTGGTGGAGTCTGGCCCTTC	120
		623	CTCGTCTGTGCTGTGTCGCAACTGTTCTCCCTGCTGGTGGAGTCTGGCCCTTC	682
QY	Db	121	CTGTCCACACCGCTGAAAGGAGCTTTCAGAGGACCGCAGGAACATCGCCCAAC	180
		683	CTGTCCACACCGCTGAAAGGAGCTTTCAGAGGACCGCAGGAACATCGCCCAAC	742
QY	Db	181	ATCATCTGTGTGTGACGAGACACAGGATGTGGAGCTGGGTTCATGCAAGTATGAAC	240
		743	ATCATCTGTGTGTGACGAGACACAGGATGTGGAGCTGGGTTCATGCAAGTATGAAC	802
QY	Db	241	AAGACCGGGGATCATGAGCAGGGCGGGGCGCACTTCATCAAGCCCTTGTGACCA	300
		803	AAGACCGGGGATCATGAGCAGGGCGGGGCGCACTTCATCAAGCCCTTGTGACCA	862
QY	Db	301	CCCATGTGTGCCCTTCAGCTCCTCCATCTCCTGCAAGTACGTCCACAACCAAC	360
		863	CCCATGTGTGCCCTTCAGCTCCTCCATCTCCTGCAAGTACGTCCACAACCAAC	922
QY	Db	361	ACCTACCAACAATAGAACTGCTCTCCCTCTGCGCCCTCTGCGAGGACAGCAGAGCCGC	420
		923	ACCTACCAACAATAGAACTGCTCTCCCTCTGCGCCCTCTGCGAGGACAGCAGAGCCGC	982
QY	Db	421	ACCTTTGCGGTGTACCTCAATAGCTGCTACCGGACAGCTTCTTCCGGGAAGTATCTT	480
		983	ACCTTTGCGGTGTACCTCAATAGCTGCTACCGGACAGCTTCTTCCGGGAAGTATCTT	1042
QY	Db	481	AATGAATACACGGCTCTTACGTGCCACCGCGCTGGAAGAGTGGGTCTCGACTCCCTTAA	540
		1043	AATGAATACACGGCTCTTACGTGCCACCGCGCTGGAAGAGTGGGTCTCGACTCCCTTAA	1102
QY	Db	541	AACCTCCCGCTTTTATAACTACACGCTGTGTGCGAAGCGGGTGAAGAAAAGCACCGCTCC	600
		1103	AACCTCCCGCTTTTATAACTACACGCTGTGTGCGAAGCGGGTGAAGAAAAGCACCGCTCC	1162
QY	Db	601	GACTACTCCAAGGATTAACCTACAGACCTCATCAATGACAGCGTGAGCTTCTTCGCG	660
		1163	GACTACTCCAAGGATTAACCTACAGACCTCATCAATGACAGCGTGAGCTTCTTCGCG	1222
QY	Db	661	AGCTCCAAGAAGATGTACCCGACAGGCGCAGTCTCATGGTTCATCAGCCATCAGCCCCC	720
		1223	AGCTCCAAGAAGATGTACCCGACAGGCGCAGTCTCATGGTTCATCAGCCATCAGCCCCC	1282
QY	Db	721	CACGGCCCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAAAGCATCTCAGCAC	780
		1283	CACGGCCCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAAAGCATCTCAGCAC	1342
QY	Db	781	ATCAGCCCGAGCTTACAACTACCGGCCCAACCCGGAACAAACACTGGATCATGCGCTACAG	840
		1343	ATCAGCCCGAGCTTACAACTACCGGCCCAACCCGGAACAAACACTGGATCATGCGCTACAG	1402
QY	Db	841	GGCCCATGAAGCCCATCCATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCG	900
		1403	GGCCCATGAAGCCCATCCATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCG	1462
QY	Db	901	ACCTCATGTCCGTGAGCGACTCCATGAGAGCGATTTCACACATCTGTTGAGACGGCG	960
		1463	ACCTCATGTCCGTGAGCGACTCCATGAGAGCGATTTCACACATCTGTTGAGACGGCG	1522
QY	Db	961	GAGCTGGACACACGCTACATCGTATACACCGCGACCGGTTACCAATCGCCAGTTT	1020
		1523	GAGCTGGACACACGCTACATCGTATACACCGCGACCGGTTACCAATCGCCAGTTT	1582
QY	Db	1021	GGCTGTGAAAGGAAATCCATGATGATGTTGACATCAGGGTCCGTTCTACGCTG	1080
		1583	GGCTGTGAAAGGAAATCCATGATGATGTTGACATCAGGGTCCGTTCTACGCTG	1642
QY	Db	1081	AGGGGCCCCAAGTGGAAGCGCGCTGTCTGAATCCCCACATCTGCTCTCAACTGACCTG	1140
		1643	AGGGGCCCCAAGTGGAAGCGCGCTGTCTGAATCCCCACATCTGCTCTCAACTGACCTG	1702
QY	Db	1141	GCSCCACAATCTGACATATTCAGGCTTGACATACCTGCGGATATGACCGGGAATCC	1200
		1703	GCSCCACAATCTGACATATTCAGGCTTGACATACCTGCGGATATGACCGGGAATCC	1762
QY	Db	1201	ATCTCAAGCTGTGACACGAGCGCGCGTGAATTCGGTTTCACTTGAAGAAAGATG	1260
		1763	ATCTCAAGCTGTGACACGAGCGCGCGTGAATTCGGTTTCACTTGAAGAAAGATG	1822
QY	Db	1261	AGGCTCTGCGGAGCTCTCTTCTGTTGAGAGAGGCAAGCTGCTACAAAGAGACAAT	1320
		1823	AGGCTCTGCGGAGCTCTCTTCTGTTGAGAGAGGCAAGCTGCTACAAAGAGACAAT	1882
QY	Db	1321	GACAAGTGGACGCCCGAGGAGGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTG	1380
		1883	GACAAGTGGACGCCCGAGGAGGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTG	1942
QY	Db	1381	TGTGAGCTGTGAGTACCAAGCGCGCTGTGACAGCTGGGACAGAGTGGAGTGTG	1440
		1943	TGTGAGCTGTGAGTACCAAGCGCGCTGTGACAGCTGGGACAGAGTGGAGTGTG	2002
QY	Db	1441	GAGGACGCCACCGGGAGGCTGAGCTGCATAGTGCAAGGGCCCAATGCGGTGGCGCGC	1500
		2003	GAGGACGCCACCGGGAGGCTGAGCTGCATAGTGCAAGGGCCCAATGCGGTGGCGCGC	2062
QY	Db	1501	AGCAGAGCCCTCTCCCAACCTCGTGCCTCAAGTACTACGGGACGGGACGAGGCGCTSCAC	1560
		2063	AGCAGAGCCCTCTCCCAACCTCGTGCCTCAAGTACTACGGGACGGGACGAGGCGCTSCAC	2122
QY	Db	1561	TGTGACAGCGGGAGCTTCAAGCTCAGCTGGCGGACGCGGGAAGAACTCTTCAGAGAG	1620
		2123	TGTGACAGCGGGAGCTTCAAGCTCAGCTGGCGGACGCGGGAAGAACTCTTCAGAGAG	2182
QY	Db	1621	AAGTACAGGCCAGCTATGTCCGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC	1680

Db 2183 AAGTACAAAGCCAGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCCGAGGTGGAC 2242  
Qy 1681 GGCAGGGGTGTACCACTAGGCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG 1740  
Db 2243 GGCAGGGGTGTACCACTAGGCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG 2302  
Qy 1741 CACTGGCCAGGGGCCCTGTAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGGGGCCCTGTAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 2362  
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCATTAAGTGTACATCGGTGCTACATC 1860  
Db 2363 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCATTAAGTGTACATCGGTGCTACATC 2422  
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGSCCTGGAAA 1920  
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGSCCTGGAAA 2482  
Qy 1921 GACCACAGCTGCACATCGACCCAGAGATTGAACCTTCGACGACCAAAATTAAGAACCTG 1980  
Db 2483 GACCACAGCTGCACATCGACCCAGAGATTGAACCTTCGACGACCAAAATTAAGAACCTG 2542  
Qy 1981 AGGGAAGTCCGAGTCACTGTAAGAAAAGCGGCCAGAGATGTGACTGTCACAAAATC 2040  
Db 2543 AGGGAAGTCCGAGTCACTGTAAGAAAAGCGGCCAGAGATGTGACTGTCACAAAATC 2602  
Qy 2041 AGTACACACCCAGCACAAGGCGGCTCAAGCAGAGGCTCCAGTCTGCATCCTTTTC 2100  
Db 2603 AGTACACACCCAGCACAAGGCGGCTCAAGCAGAGGCTCCAGTCTGCATCCTTTTC 2662  
Qy 2101 AGGAAGGGCTGCAAGAGAGAGCAAGTGTGCTGTGCGGAGCAGAGCGCAGAGAG 2160  
Db 2663 AGGAAGGGCTGCAAGAGAGAGCAAGTGTGCTGTGCGGAGCAGAGCGCAGAGAG 2722  
Qy 2161 AAATCCGCAAGCTGTCAAGCGCCTGCAAGCAACAGACAGTGCAGATGCGCAGGCTC 2220  
Db 2723 AAATCCGCAAGCTGTCAAGCGCCTGCAAGCAACAGACAGTGCAGATGCGCAGGCTC 2782  
Qy 2221 ACGTGCTTCCACGACACCGACTGCGACAGCGGCTTTCTGGACACTGGGGCT 2280  
Db 2783 ACGTGCTTCCACGACACCGACTGCGACAGCGGCTTTCTGGACACTGGGGCT 2842  
Qy 2281 TTCTGTGCTGCGACAGCGCCCAACATTAACACGCTACTGTGTCATGAGGACCATCAATGAG 2340  
Db 2843 TTCTGTGCTGCGACAGCGCCCAACATTAACACGCTACTGTGTCATGAGGACCATCAATGAG 2902  
Qy 2341 ACTCAAAATTCCTCTCTGTGAAATTTGCAACTGGCTTCCCTAGAGTACTTTGATCTCAAC 2400  
Db 2903 ACTCAAAATTCCTCTCTGTGAAATTTGCAACTGGCTTCCCTAGAGTACTTTGATCTCAAC 2962  
Qy 2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGACAGGATGCTCTCAACCCAG 2460  
Db 2963 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGACAGGATGCTCTCAACCCAG 3022  
Qy 2461 CTACACGTCAGCTCACTGGAGCTGAGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCCGG 2520  
Db 3023 CTACACGTCAGCTCACTGGAGCTGAGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCCGG 3082  
Qy 2521 ACTCGAACAATGGACCTGGACTTAAAGTGGAGGAGCTATGAGCAATACAGGCAATTT 2580  
Db 3083 ACTCGAACAATGGACCTGG-----GATGGGGAAGCTATGAGCAATACAGGCAATTT 3133  
Qy 2581 CAGCGTCGAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 2640  
Db 3134 CAGCGTCGAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 3193  
Qy 2641 GRAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGAACCTTCCAAAAAATAGAGGCAATCAC 2700  
Db 3194 GAAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGAACCTTCCAAAAAATAGAGGCAATCAC 3253  
Qy 2701 TGACTGCACAGGCAATGAAAAACCATGTGGTGTATTTCCAGAGACCTGTGTATTGGCC 2760

Db 3254 TGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313  
Qy 2761 AGGAGGCTGTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTGTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373  
Qy 2821 CAGCAGGACGAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 2880  
Db 3374 CAGCAGGACGAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 3433  
Qy 2881 TACTCTACCGCTCCACAAATGCAATTTTCTGATCAAAAGTGCACACCTTAACCCCTCC 2940  
Db 3434 TACTCTACCGCTCCACAAATGCAATTTTCTGATCAAAAGTGCACACCTTAACCCCTCC 3493  
Qy 2941 CCAGAGCTCAAAAGGAAAAACGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3000  
Db 3494 CCAGAGCTCAAAAGGAAAAACGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3553  
Qy 3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3060  
Db 3554 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3613  
Qy 3061 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGGAACCTAAAGAGCAGCAGAGCAACGT 3120  
Db 3614 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGGAACCTAAAGAGCAGCAGAGCAACGT 3673  
Qy 3121 GGAGAGGCTGAAAAACAGTGCAGAGACCTTTTGAATGAGTCAATGAGTCAACAAAGAGATGA 3180  
Db 3674 GGAGAGGCTGAAAAACAGTGCAGAGACCTTTTGAATGAGTCAATGAGTCAACAAAGAGATGA 3733  
Qy 3181 CATTTACCTAGCCTATAAACCCCTGTTGCTCTGAGAGAACTGCCCTTCATTTGATATAT 3240  
Db 3734 CATTTACCTAGCCTATAAACCCCTGTTGCTCTGAGAGAACTGCCCTTCATTTGATATAT 3793  
Qy 3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAAACCTTAATAAGAAATCCCAA 3300  
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAAACCTTAATAAGAAATCCCAA 3853  
Qy 3301 TTTTTCAGAGTGGTGTGTCATTAACCGCTCTGTGGCCAGTGTAAAGAAAAA 3353  
Db 3854 TTTTTCAGAGTGGTGTGTCATTAACCGCTCTGTGGCCAGTGTAAAGAAAAA 3906

Search completed: February 14, 2004, 22:52:51  
Job time : 1148 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 14, 2004, 11:41:13 ; Search time 11897 Seconds  
(without alignments)  
11564.182 Million cell updates/sec

Title: US-10-025-966A-4  
Perfect score: 3363  
Sequence: 1 gactccgcatcccaaaaga.....aaaagaaaaaa 3363

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_hgt:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sta:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sta:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_hgt\_hum:\*
- 31: em\_hgt\_inv:\*
- 32: em\_hgt\_other:\*
- 33: em\_hgt\_mus:\*
- 34: em\_hgt\_pln:\*
- 35: em\_hgt\_rod:\*
- 36: em\_hgt\_man:\*
- 37: em\_hgt\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgt\_hum:\*
- 40: em\_hgt\_mus:\*
- 41: em\_hgt\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3363	100.0	3363	6	AX573182 Sequence
2	3347.2	99.5	4397	9	AB033073 Homo sapi
3	3346.8	99.5	4279	9	AY101176 Homo sapi
4	3346.8	99.5	4286	6	AX573191 Sequence
5	3344.8	99.5	3348	6	AX211371 Sequence
6	3323.4	98.8	3906	6	AR252433 Sequence
7	3323.4	98.8	3906	6	AX403196 Sequence
8	3323.4	98.8	3906	6	AX454488 Sequence
9	3323.4	98.8	3906	6	AX463886 Sequence
10	3323.4	98.8	3906	6	AX490966 Sequence
11	3223.2	95.8	3513	6	AX470095 Sequence
12	2654.8	78.9	3289	6	AX573192 Sequence
13	2613	77.7	2616	6	AX573183 Sequence
14	2605.2	77.5	2610	6	AX470097 Sequence
15	2294.8	68.2	2559	6	BD127910 Primer fo
16	2294.8	68.2	2559	9	AK074483 Homo sapi
17	2174.8	64.7	2178	6	BD127849 Primer fo
18	2174.8	64.7	2178	9	AK075035 Homo sapi
19	2095.2	62.3	3603	10	AY101177 Mus muscu
20	2095.2	62.3	3613	6	AX573194 Sequence
21	2095.2	62.3	3766	6	AX573188 Sequence
22	2058.2	61.2	2628	6	AX573189 Sequence
23	2058.2	61.2	2628	6	AX573195 Sequence
24	1727.4	51.4	1729	9	HS049G161 Novel hum
25	1625	48.3	1661	9	BC020962 Homo sapi
26	1151	34.2	3688	5	BC045403 Danio rer
27	1080.4	32.1	4088	10	AF230072 Rattus no
28	1070.6	31.8	5769	5	AF410802 Coturnix
29	1062.4	31.6	4623	10	BC034547 Mus muscu
30	1062.4	31.6	4623	10	BC049276 Mus muscu
31	1062.4	31.6	4671	10	AY101178 Mus muscu
32	1040.4	30.9	2616	6	AX206971 Sequence
33	1040.4	30.9	2616	6	AX573180 Sequence
34	1040.4	30.9	4321	6	AX206963 Sequence
35	1040.4	30.9	4378	6	AX188060 Sequence
36	1040.4	30.9	4378	6	AX188065 Sequence
37	1040.4	30.9	4378	6	AX188144 Sequence
38	1040.4	30.9	4378	6	AX188145 Sequence
39	1040.4	30.9	4378	6	AX188147 Sequence
40	1040.4	30.9	4378	6	AX188329 Sequence
41	1040.4	30.9	4378	6	AX188343 Sequence
42	1040.4	30.9	4669	9	AY101175 Homo sapi
43	1040.4	30.9	4961	6	AX463510 Sequence
44	1040.4	30.9	5244	6	AX573179 Sequence
45	1038.8	30.9	5699	9	AF545571 Homo sapi

ALIGNMENTS

RESULT 1	AX573182	AX573182	3363 bp	DNA	linear	PAT 29-NOV-2002
LOCUS	Sequence 4 from Patent WO02059327.					
DEFINITION	AX573182					
ACCESSION	AX573182.1	GI:26005103				
VERSION						
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE 1						
AUTHORS	Rosen,S.D., Hemmerich,S., Tomita,M. and Palmeri,D.					
TITLE	Sulfatases and methods of use thereof					
JOURNAL	Patent: WO 02059327-A 4 01-AUG-2002;					









1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGACGGCCTCGAAA 1920  
2146 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGACGGCCTCGAAA 2205  
1921 GACCACAAAGCTGCACATCGACACGAGATTCAACCTCTGACAGACAAATTAAGAACCTG 1980  
2206 GACCACAAAGCTGCACATCGACACGAGATTCAACCTCTGACAGACAAATTAAGAACCTG 2265  
1981 AGGAAAGTCCGAGTCACTCTGAAGAAAAGCGGCAGAGAAATGTGACTGTCAAAAATC 2040  
2266 AGGAAAGTCCGAGTCACTCTGAAGAAAAGCGGCAGAGAAATGTGACTGTCAAAAATC 2325  
2041 AGCTACACACCCGACCAAGGCGCCTCAAGCAGAGAGCTCCAGTCTGCATCTCTTC 2100  
2326 AGCTACACACCCGACCAAGGCGCCTCAAGCAGAGAGCTCCAGTCTGCATCTCTTC 2385  
2101 AGGAAGGCGCTCGAAGAGAGGACAGAGTGTGGCTGTGGCGAGCAGAGCGCAAGAAG 2160  
2386 AGGAAGGCGCTCGAAGAGAGGACAGAGTGTGGCTGTGGCGAGCAGAGCGCAAGAAG 2445  
2161 AAATCTCGCAAGCTGTCAAGCGCTCTGAGAAACAAACACAGTGCAGATGCGAGCGCTC 2220  
2446 AAATCTCGCAAGCTGTCAAGCGCTCTGAGAAACAAACACAGTGCAGATGCGAGCGCTC 2505  
2221 ACGTGCTTACCCACGACCAACAGCAGTCTGCGAGAGCGCGCTTCTGAGACACTGGGCGCT 2280  
2506 ACGTGCTTACCCACGACCAACAGCAGTCTGCGAGAGCGCGCTTCTGAGACACTGGGCGCT 2565  
2281 TTCTGTGCTGCTGACAGCGCCCAACAAATAACAGTGTGCTGATGAGGACCAATCAATGAG 2340  
2566 TTCTGTGCTGCTGACAGCGCCCAACAAATAACAGTGTGCTGATGAGGACCAATCAATGAG 2625  
2341 ACTCACAATTTCTCTCTGTAATTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
2626 ACTCACAATTTCTCTCTGTAATTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2685  
2401 ACAGACCCCTACAGCTGATGAATCAGTGAACACACTGGACAGCGGATGCTCTCAACCCAG 2460  
2686 ACAGACCCCTACAGCTGATGAATCAGTGAACACACTGGACAGCGGATGCTCTCAACCCAG 2745  
2461 CTACACGTACAGCTCATGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2520  
2746 CTACACGTACAGCTCATGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2805  
2521 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGCGAGTTT 2580  
2806 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGCGAGTTT 2865  
2581 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATACATGCGACAACTGTGG 2640  
2866 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATACATGCGACAACTGTGG 2925  
2641 GAAGCTGGGAGGTTAAGAACACACAGAGTGGAGCTCCCAAAACATAGAGCATCACC 2700  
2926 GAAGCTGGGAGGTTAAGAACACACAGAGTGGAGCTCCCAAAACATAGAGCATCACC 2985  
2701 TGACTGCACAGGCAATGAAGAACCACTGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760  
2986 TGACTGCACAGGCAATGAAGAACCACTGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 3045  
2761 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATCTGGAGGATAC 2820  
3046 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATCTGGAGGATAC 3105  
2821 CAGCAGGAGCAGAGTAATCTTCAAGAGTCCATTTTGGCCCTGCTTTTGGCTTGGATTA 2880  
3106 CAGCAGGAGCAGAGTAATCTTCAAGAGTCCATTTTGGCCCTGCTTTTGGCTTGGATTA 3165  
2881 TACCTCAGCAGCTGCAAAAATGATTTTTCGTATCAAAAAGTCAACACTAACCTTCCC 2940  
3166 TACCTCAGCAGCTGCAAAAATGATTTTTCGTATCAAAAAGTCAACACTAACCTTCCC 3225  
2941 CCAGAGCTCAGAAAGGAAACCGGAGAGCGGAGAGAGATTTCTCTTGGAAATTTCT 3000

3226 CCAGAAGCTCAAAAGGAAAACGGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3285  
3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGACAGTCTCTGTTCTAAA 3060  
3286 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGACAGTCTCTGTTCTAAA 3345  
3061 TCTCTTATCTTTTGGTTTGTCAAAAGAGGAACTAAGAGAGGAGGACAGAGGCAACGT 3120  
3346 TCTCTTATCTTTTGGTTTGTCAAAAGAGGAACTAAGAGAGGAGGACAGAGGCAACGT 3405  
3121 GGAGAGGCTGAAACAGTGCAGAGACGTTTCACAAATGAGTCAAGTACAGCAAAAGAGATGA 3180  
3406 GGAGAGGCTGAAACAGTGCAGAGACGTTTCACAAATGAGTCAAGTACAGCAAAAGAGATGA 3465  
3181 CATTTACCTAGCAGTATAAACCTCGTTGCTCTCAAGAAACTGCTTCAATTTGATATAT 3240  
3466 CATTTACCTAGCAGTATAAACCTCGTTGCTCTCAAGAAACTGCTTCAATTTGATATAT 3525  
3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACTTAAGAAATCCCAA 3300  
3526 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACTTAAGAAATCCCAA 3585  
3301 TTTTCAGGAGTGGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3352  
3586 TTTTCAGGAGTGGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3637  
RESULT 3  
AY101176  
LOCUS  
DEFINITION Homo sapiens extracellular sulfatase SULF-2 mRNA linear PRI 21-DEC-2002  
ACCESSION AY101176  
VERSION AY101176.1 GI:27356933  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Morimoto-Tomita,M., Uchimura,K., Werb,Z., Hemmerich,S. and Rosen,S.D.  
TITLE Cloning and Characterization of Two Extracellular Heparin-degrading Endosulfatases in Mice and Humans  
J. Biol. Chem. 277 (51), 49175-49185 (2002)  
JOURNAL  
PUBMED 12368295  
REFERENCE  
AUTHORS Morimoto-Tomita,M., Werb,Z., Hemmerich,S. and Rosen,S.D.  
TITLE Direct Submission  
Submitted (30-APR-2002) Biology, Thios Pharmaceuticals Inc., 747 Fifty Second Street, Oakland, CA 94609, USA  
JOURNAL  
FEATURES  
Location/Qualifiers  
1. 4279  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q13.12"  
/tissue\_type="lung"  
670..3262  
/codon\_start=1  
/product="extracellular sulfatase SULF-2"  
/protein\_id="AA076861.1"  
/db\_xref="GI:27356934"  
CDS  
1  
translation="MGPPSLVLCILSATVSLGGSAFLSHRLKGRFQDRNRNP  
NIIIVTDQDELGSMVKNKTRIMBOGGAHFINAVTTPMCCPSRSSLTGKYVH  
NNTVNNCSPSWQAQHESTRFVYVNSTGYRTAFYFGKYLNEVNGSVYPPQWKEW  
VGLKASRPVNYTLCRNGVEKESDYKDYLTDLITNDSVSFFRTSKMYPHRPLVM  
VISHAPGPDSDAPQSRLEFPNASQHTPSYNYAPAPDKHWNRYTGPMPKPIHEFT  
NMLQRRLTLMSVDDSNETIYNMLVETGELDNTIYITADHGHYIQFGLVKGKSWP  
YEFIRVFFYVRGNFENCLNPHIVLNIDIAPIIDAGLDIPADMDGKSLKLLDT  
ERPYNRFLKKKRWDRDSFLVERGKLHKRDNDKVDQAQENFLPKYQVRKDLQRAE  
YOTACEQLGQXQWQCVEDATGKLKHKCKPWLKGRSRLSNLVPKTYQGSEACTCDS

GDYKLSLAGRRKKLKKYKASIVRSRSIRSAIEVDGRVYHVLGIDAAQPRNLTKRH  
WFGAEDQDDXGGPSFGYGPDIYSAANPIKVTIRCVILENDTVQCDLDLYKSLQAW  
KDHKLHI DEIEITLONKI NLRREVRLHKKRPEEDCHKI SYHTOHKRLKRGSSSL  
HPRKLGLOSKDKVLLREOKRKKLRLKLLRLONNDTCSMPGLTCTFDHNOHWOTAPF  
WTLGPFCATCSANNITWCMFTINETHNLFCEPATGLEVEDLNTDPOYOLMNAVNL  
DRDLNLQHLVQMLRSLCKYKQCNPRTRNDMLGLKDGGSYEQYRQFORRWPEKRP  
SRKSLGQLMEGWEH\*

polyA signal 3944.3948  
BASE COUNT 1186 a 1143 c 1105 g 845 t  
ORIGIN

Query Match		99.5%;	Score 3346.8;	DB 9;	Length 4279;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 3359;		Conservative	0;	Mismatches	2;
				Indels	1;
				Gaps	1;
QY	1	GACTCCCGCATCCCAAAGAAGCACCAGATCAGCAAAAAGAGATGGCCCCCGGAGC	60		
DB	625	GACTCCCGCATCCCAAAGAAGCACCAGATCAGCAAAAAGAGATGGCCCCCGGAGC	684		
QY	61	CTCGTGTGCTGTGCTGTCCGCACTGTCTTCCCTGCTGGTGGAGCTCGGCTTC	120		
DB	685	CTCGTGTGCTGTGCTGTCCGCACTGTCTTCCCTGCTGGTGGAGCTCGGCTTC	744		
QY	121	CTGTCGCAACACCGCTGAAAGAGCAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC	180		
DB	745	CTGTCGCAACACCGCTGAAAGAGCAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC	804		
QY	181	ATCATCTGTGTGTCAGGACGACAGATGTGGAGCTGGTTCATGAGGTGATGAAC	240		
DB	805	ATCATCTGTGTGTCAGGACGACAGATGTGGAGCTGGTTCATGAGGTGATGAAC	864		
QY	241	AAGACCCGGCGCATCATGAGCAGGCGGGGGCGCACTTCATCAAGCCCTTCGTGACCA	300		
DB	865	AAGACCCGGCGCATCATGAGCAGGCGGGGGCGCACTTCATCAAGCCCTTCGTGACCA	924		
QY	301	CCCATGTGTGCCCTTCACGCTCTCCATPCCTCACTGCGCAAGTACGTCCCAACCAAC	360		
DB	925	CCCATGTGTGCCCTTCACGCTCTCCATPCCTCACTGCGCAAGTACGTCCCAACCAAC	984		
QY	361	ACCTACACCAACATGAGAACTCTCTCGCCCTCTGCGAGGACACACAGAGCGCG	420		
DB	985	ACCTACACCAACATGAGAACTCTCTCGCCCTCTGCGAGGACACACAGAGCGCG	1044		
QY	421	ACCTTTGCGGTGATCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT	480		
DB	1045	ACCTTTGCGGTGATCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT	1104		
QY	481	AATGAATACAGCGCTCTACGTGCGACCCGGCTGGAGAGTGGTCTGGACTCTTAA	540		
DB	1105	AATGAATACAGCGCTCTACGTGCGACCCGGCTGGAGAGTGGTCTGGACTCTTAA	1164		
QY	541	AACCTCCGCTTTTATACTACAGCTGTCTCGGAACGGGTGAAAGAAAGCAGCGCTCC	600		
DB	1165	AACCTCCGCTTTTATACTACAGCTGTCTCGGAACGGGTGAAAGAAAGCAGCGCTCC	1224		
QY	601	GACTATCCAGGATTAATCTCAGACACTTCATCAACATGACAGCGTGAAGTCTTCGCG	660		
DB	1225	GACTATCCAGGATTAATCTCAGACACTTCATCAACATGACAGCGTGAAGTCTTCGCG	1284		
QY	661	ACGTCCAGAGATGATACCGGACAGCGCAGTCTCATGCTCATCAGCATGACGCCGCC	720		
DB	1285	ACGTCCAGAGATGATACCGGACAGCGCAGTCTCATGCTCATCAGCATGACGCCGCC	1344		
QY	721	CACGGCCCTGAGGATTCAGCCCCACAAATATTCACGCTCTTCCCAAAGCATCTCAGCAC	780		
DB	1345	CACGGCCCTGAGGATTCAGCCCCACAAATATTCACGCTCTTCCCAAAGCATCTCAGCAC	1404		
QY	781	ATCACGCGAGCTTAACTACGCGCCACACCGGACCAACATGATCATGCGCTACAGC	840		
DB	1405	ATCACGCGAGCTTAACTACGCGCCACACCGGACCAACATGATCATGCGCTACAGC	1464		
QY	841	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGAG	900		

1465	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGAAGCGCTTGAC	1524
901	ACCTCATGTGCGGTGACAGCTCCATGAGAGAGATTTACAAATGCTGCTGAGACGGGC	960
1525	ACCTCATGTGCGGTGACAGCTCCATGAGAGAGATTTACAAATGCTGCTGAGACGGGC	1584
961	GAGCTGGAACAAACGTGATCATGCTATACACCGCGGACACGGTTACACATCGGCCAGTTT	1020
1585	GAGCTGGAACAAACGTGATCATGCTATACACCGCGGACACGGTTACACATCGGCCAGTTT	1644
1021	GGCTGTGTAAGGGAATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1080
1645	GGCTGTGTAAGGGAATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1704
1081	AGGGGCCCCAACGTGGAAGCGGCTCTGTGAATCCCAACATGCTCTCAACATGACCTG	1140
1705	AGGGGCCCCAACGTGGAAGCGGCTCTGTGAATCCCAACATGCTCTCAACATGACCTG	1764
1141	GGCCCCACCATCTGTCACATTCGAGGCTGGACATACCTGCGGATATGACGGGAATCC	1200
1765	GGCCCCACCATCTGTCACATTCGAGGCTGGACATACCTGCGGATATGACGGGAATCC	1824
1201	ATCCTCAAGCTGCTGACACCGGAGCGGCTGAATCGGTTTCACTTGAATAAGAGATG	1260
1825	ATCCTCAAGCTGCTGACACCGGAGCGGCTGAATCGGTTTCACTTGAATAAGAGATG	1884
1261	AGGCTCTGGCGGACCTCTTCTGTTGGAGAGGCAAGCTGCTACACAGAGAGACAAT	1320
1885	AGGCTCTGGCGGACCTCTTCTTGGTGGAGAGGCAAGCTGCTACACAGAGAGACAAT	1944
1321	GACAAAGTGACGCCCCAGGAGGAGAACTTTCTGCCCAAGTACCGAGCTGTGAAGACCTG	1380
1945	GACAAAGTGACGCCCCAGGAGGAGAACTTTCTGCCCAAGTACCGAGCTGTGAAGACCTG	2004
1381	TGTCAGGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG	1440
2005	TGTCAGGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAACTGGCAGTGTGTG	2064
1441	GAGAGCGCCACAGGGGAAGCTGAAGCTGCATAGTCAAGGCGCCCATGGGCTGGGCGGC	1500
2065	GAGAGCGCCACAGGGGAAGCTGAAGCTGCATAGTCAAGGCGCCCATGGGCTGGGCGGC	2124
1501	AGCAGAGCCCTCTCCAACTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	1560
2125	AGCAGAGCCCTCTCCAACTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	2184
1561	TGTGACAGCGGGGACCTACAGCTCAGCTGCGCGGACCGCGGAAATACTCTTCAAGAAG	1620
2185	TGTGACAGCGGGGACCTACAGCTCAGCTGCGCGGACCGCGGAAATACTCTTCAAGAAG	2244
1621	AAGTACAAAGGCCAAGCTATGTCCGAGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC	1680
2245	AAGTACAAAGGCCAAGCTATGTCCGAGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC	2304
1681	GGCAGGCTGACCACTGAGGCTGGTGTGATGCGCGCCAGCCCGGAAACCTCCACAGCGG	1740
2305	GGCAGGCTGACCACTGAGGCTGGTGTGATGCGCGCCAGCCCGGAAACCTCCACAGCGG	2364
1741	CATCGGCCAGGGGCCCTCTGAGGACCAAGATGACAAAGATGTTGGGGACCTTCAGTGGCACT	1800
2365	CATCGGCCAGGGGCCCTCTGAGGACCAAGATGACAAAGATGTTGGGGACCTTCAGTGGCACT	2424
1801	GAGAGGCTTCCCGACTATCTCAGCGGCCAACCCTATTAAGTGAACATCGGTGCTACATC	1860
2425	GAGAGGCTTCCCGACTATCTCAGCGGCCAACCCTATTAAGTGAACATCGGTGCTACATC	2484
1861	CTAGAGAACACACAGCTCCAGTGTGACCTGTACAAGTCTCCCTGACGAGCCTCGAAA	1920
2485	CTAGAGAACACACAGCTCCAGTGTGACCTGTACAAGTCTCCCTGACGAGCCTCGAAA	2544
1921	GACCACAGCTGCATCGACCAAGATGTGAAAACCTTGACAGAACAAAATTAAGAACCTG	1980
2545	GACCACAGCTGCATCGACCAAGATGTGAAAACCTTGACAGAACAAAATTAAGAACCTG	2604

```
QY 1981 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCAGAGAAATGTGACTGTGCACAAAATC 2040
Db 2605 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCAGAGAAATGTGACTGTGCACAAAATC 2664
QY 2041 AGCTACCAACCCAGCACAAGAGCGCCCTCAAGCACAGAGCTCCAGTCTGCATCCTTTC 2100
Db 2665 AGCTACCAACCCAGCACAAGAGCGCCCTCAAGCACAGAGCTCCAGTCTGCATCCTTTC 2724
QY 2101 AGGAAGGSCCTGCAAGAGAGGACAAGGTGTGGCTGTTCGGGAGCAGAAGCGCAAGAAG 2160
Db 2725 AGGAAGGSCCTGCAAGAGAGGACAAGGTGTGGCTGTTCGGGAGCAGAAGCGCAAGAAG 2784
QY 2161 AAACTCCCAAGCTCTCAAGCGCTCGAGAAACAAACGACAGTGCAGCATGCCAGGCTTC 2220
Db 2785 AAACTCCCAAGCTCTCAAGCGCTCGAGAAACAAACGACAGTGCAGCATGCCAGGCTTC 2844
QY 2221 ACGTGTCTTCAACCAGCAACACGACCTGGCAGACGGCGCTTCTTGACACATGGGGCCT 2280
Db 2845 ACGTGTCTTCAACCAGCAACACGACCTGGCAGACGGCGCTTCTTGACACATGGGGCCT 2904
QY 2281 TTCTGTGCTGCACAGCGCCAAACAATACACAGTACTGTGTCATGAGGACCATCAATGAG 2340
Db 2905 TTCTGTGCTGCACAGCGCCAAACAATACACAGTACTGTGTCATGAGGACCATCAATGAG 2964
QY 2341 ACTCACAATTTCTCTCTGCTGTAATTTGCACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
Db 2965 ACTCACAATTTCTCTCTGCTGTAATTTGCACTGGCTTCTTAGAGTACTTTGATCTCAAC 3024
QY 2401 ACAGACCCCTACCGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTTCAACAG 2460
Db 3025 ACAGACCCCTACCGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTTCAACAG 3084
QY 2461 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTACAAGCAGTGTAAACCCCGG 2520
Db 3085 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTACAAGCAGTGTAAACCCCGG 3144
QY 2521 ACTCGAAACATGGACCTTGGGACTTAAAGATGAGCAAGCTATGAGCAATACAGCAGTTT 2580
Db 3145 ACTCGAAACATGGACCTTGGGACTTAAAGATGAGCAAGCTATGAGCAATACAGCAGTTT 3204
QY 2581 CAGCTGCAAGGTGCGCAGAAATGAAGAGACTTCTTCCAAATCACTGGACAACTGTGG 2640
Db 3205 CAGCTGCAAGGTGCGCAGAAATGAAGAGACTTCTTCCAAATCACTGGACAACTGTGG 3264
QY 2641 GAAGGCTGGAGGTTAAGAACACACAGAGTGAGCTCCCAAAACATAGAGCATCAAC 2700
Db 3265 GAAGGCTGGAGGTTAAGAACACACAGAGTGAGCTCCCAAAACATAGAGCATCAAC 3324
QY 2701 TGACTGCAAGGCAATGAAACCAATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3325 TGACTGCAAGGCAATGAAACCAATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3384
QY 2761 AGGAGGCTGAGAAAGCAGACGCACTCTCAGTCAACATGA CAGATTCGAGAGGATAAC 2820
Db 3385 AGGAGGCTGAGAAAGCAGACGCACTCTCAGTCAACATGA CAGATTCGAGAGGATAAC 3444
QY 2821 CAGCAGGACGAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGCCTTTGGATTA 2880
Db 3445 CAGCAGGACGAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGCCTTTGGATTA 3504
QY 2881 TACTCTACCAAGTGCACAAAATGATTTTTCGATCAAAAAGTCAACCACTAACCCCTCC 2940
Db 3505 TACTCTACCAAGTGCACAAAATGATTTTTCGATCAAAAAGTCAACCACTAACCCCTCC 3564
QY 2941 CCAGAAGCTCAAGAAGAAACGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3565 CCAGAAGCTCAAGAAGAAACGAGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3624
QY 3001 CCCAAGGCGAAAGTCAATTGGAATTTTAAATCA TAGGGGAAAAGCAGTCTGTCTTAA 3060
Db 3625 CCCAAGGCGAAAGTCAATTGGAATTTTAAATCA TAGGGGAAAAGCAGTCTGTCTTAA 3684
```

```
QY 3061 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAACTAAGAGCAGGACAGAGCAAGCT 3120
Db 3685 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAACTAAGAGCAGGACAGAGCAAGCT 3744
QY 3121 GGAGAGGCTGAAAAACAGTGCAGAGACGTTTGACATAGTCTAGTAGCACAAGAGATGA 3180
Db 3745 GGAGAGGCTGAAAAACAGTGCAGAGACGTTTGACATAGTCTAGTAGCACAAGAGATGA 3804
QY 3181 CATTTACCTTAGCACTATAAACCCCTGTTGCCCTCTCAAGAAACTGCTTCATTGTATAT 3240
Db 3805 CATTTACCTTAGCA-TATAAACCCCTGTTGCCCTCTCAAGAAACTGCTTCATTGTATAT 3863
QY 3241 GTGACTATTTTACATGTAATCAACATGGGAACTTTTATAGGGGAACTTAATAAGAAATCCCAA 3300
Db 3864 GTGACTATTTTACATGTAATCAACATGGGAACTTTTATAGGGGAACTTAATAAGAAATCCCAA 3923
QY 3301 TTTTTCAGGAGTGGTGTGTCATATAAACCTCTGTGGCAGTGTAAAGAAAGAAAAAAA 3360
Db 3924 TTTTTCAGGAGTGGTGTGTCATATAAACCTCTGTGGCAGTGTAAAGAAAGAAAAAAA 3983
QY 3361 AA 3362
Db 3984 AA 3985

RESULT 4
AX573191 4296 bp DNA linear PAT 29-NOV-2002
LOCUS Sequence 13 from Patent WO02059327.
DEFINITION AX573191
ACCESSION AX573191
VERSION AX573191.1 GI:26005109
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Rosen,S.D., Hemmerich,S., Tomita,M. and Palmeri,D.
TITLE Sulfatases and methods of use thereof
JOURNAL Patent: WO 02059327-A 13 01-AUG-2002;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US)
FEATURES
Location/Qualifiers
1..4286
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1188 a 1145 c 1108 g 845 t
ORIGIN

Query Match 99.5%; Score 3346.8; DB 6; Length 4286;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GACTCCCGCATCCCAAGAGAGCCAGATCAGCAAAAAAGAGATGGGCCCCCGGAGC 60
Db 632 GACTCCCGCATCCCAAGAGAGCCAGATCAGCAAAAAAGAGATGGGCCCCCGGAGC 691
QY 61 CTCGTGCTGTGCTCTCTCGCAACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCCTTC 120
Db 692 CTCGTGCTGTGCTCTCTCGCAACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCCTTC 751
QY 121 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAAATCCGCCCCCAAC 180
Db 752 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAAATCCGCCCCCAAC 811
QY 181 ATCATCTGTGCTGACGAGCGACGATGTGGAGCTGGGTTCCATGCAGGTGATGAAC 240
Db 812 ATCATCTGTGCTGACGAGCGACGATGTGGAGCTGGGTTCCATGCAGGTGATGAAC 871
QY 241 AAGACCCCGCGCATCATGAGCAGCGCGGGGCGCACTTTCATCAACGCTTCTGTCACACA 300
Db 872 AAGACCCCGCGCATCATGAGCAGCGCGGGGCGCACTTTCATCAACGCTTCTGTCACACA 931
```



301 QY CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGGCAAGTAGTCCACAACCAAC 360  
932 Db CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGGCAAGTAGTCCACAACCAAC 991  
361 QY ACCTACACCAACAATGAGAACCTCTCTCGCCCTCTCGCAGGACACAGAGAGCGC 420  
992 Db ACCTACACCAACAATGAGAACCTCTCTCGCCCTCTCGCAGGACACAGAGAGCGC 1051  
421 QY ACCTTTCGCGTGATCCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480  
1052 Db ACCTTTCGCGTGATCCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 1111  
481 QY ATGTAATACACGGCTCTACGTGTCACCCGCTCGAAGAGTGGTGGACTCTCTTAAA 540  
1112 Db ATGTAATACACGGCTCTACGTGTCACCCGCTCGAAGAGTGGTGGACTCTCTTAAA 1171  
541 QY AACTCCCGCTTTTAACTACACGCTGTGTGCGAACGGGTGAAAGAAAGCAGCGCTCC 600  
1172 Db AACTCCCGCTTTTAACTACACGCTGTGTGCGAACGGGTGAAAGAGAGCAGCGCTCC 1231  
601 QY GACTACTCAAGATTAACCTACAGACCTCATCAACAATGACAGGCTGAGCTTCTTCGC 660  
1232 Db GACTACTCAAGATTAACCTACAGACCTCATCAACAATGACAGGCTGAGCTTCTTCGC 1291  
661 QY AGTCTCAAGAGATGATACCCGACAGGCGCAGTCTCATGCTCATCAGCCATGACGCCCC 720  
1292 Db AGTCTCAAGAGATGATACCCGACAGGCGCAGTCTCATGCTCATCAGCCATGACGCCCC 1351  
721 QY CAGGCGCTGAGGATTCAGCCCAACAATTCACGCTCTTTCCAAAGCATCTCAGCAC 780  
1352 Db CAGGCGCTGAGGATTCAGCCCAACAATTCACGCTCTTTCCAAAGCATCTCAGCAC 1411  
781 QY ATCAGCGCAGCTACAACTACCGGCCCAACCGGACAAACACTGGATCATGGCTACAG 840  
1412 Db ATCAGCGCAGCTACAACTACCGGCCCAACCGGACAAACACTGGATCATGGCTACAG 1471  
841 QY GGGCCCATGAAGCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCG 900  
1472 Db GGGCCCATGAAGCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCG 1531  
901 QY ACCCTCATGTCCGTGACAGCTCCATGGAGAGCATTTACAACTGCTGGTAGACGGC 960  
1532 Db ACCCTCATGTCCGTGACAGCTCCATGGAGAGCATTTACAACTGCTGGTAGACGGC 1591  
961 QY GAGCTGGACAAACAGTACATGCTATACCGCGGACCAAGCTTACACATCGGCCAGTTT 1020  
1592 Db GAGCTGGACAAACAGTACATGCTATACCGCGGACCAAGCTTACACATCGGCCAGTTT 1651  
1021 QY GGCTTGGTGAAGGAAATCCATGTCATATGATGTTGACATCAGGGTCCCGTTTACCTG 1080  
1652 Db GGCTTGGTGAAGGAAATCCATGTCATATGATGTTGACATCAGGGTCCCGTTTACCTG 1711  
1081 QY AGGGCCCCAAGCTGAAGCGGCTGTCTGAATCCCACTGCTCCTCAACATTGACCTG 1140  
1712 Db AGGGCCCCAAGCTGAAGCGGCTGTCTGAATCCCACTGCTCCTCAACATTGACCTG 1771  
1141 QY GCCCCCAACCATCTGGACATTCAGGCTCGGACATACCTGCGGATATGACGGGAAATCC 1200  
1772 Db GCCCCCAACCATCTGGACATTCAGGCTCGGACATACCTGCGGATATGACGGGAAATCC 1831  
1201 QY ATCTCAAGCTGTGACACAGGAGCGGCTGTGAATCGGTTTCACTTGAAGAGAGATG 1260  
1832 Db ATCTCAAGCTGTGACACAGGAGCGGCTGTGAATCGGTTTCACTTGAAGAGAGATG 1891  
1261 QY AGGGCTGTGGCGGACTCTTCTTGTGGAGAGAGGCAAGCTGCTACACAAGAGACCAAT 1320  
1892 Db AGGGCTGTGGCGGACTCTTCTTGTGGAGAGAGGCAAGCTGCTACACAAGAGACCAAT 1951  
1321 QY GACAAGTGACGCCAGGAGAGAACTTTCTGCCCCAAGTACACAGGCTGTGAAGACCTG 1380  
1952 Db GACAAGTGACGCCAGGAGAGAACTTTCTGCCCCAAGTACACAGGCTGTGAAGACCTG 2011  
1381 QY TGTACGGCTGCTGAGTACACAGCGGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440

2012 Db TGTACGGCTGCTGAGTACACAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2071  
1441 QY GAGGACGCCACCGGGGAAGCTGAAGCTGCATAAGTGCAGAGGCCCCCATGCGGTGGCGGC 1500  
2072 Db GAGGACGCCACCGGGGAAGCTGAAGCTGCATAAGTGCAGAGGCCCCCATGCGGTGGCGGC 2131  
1501 QY AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGACGGGACGAGCGCTGCACC 1560  
2132 Db AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGACGGGACGAGCGCTGCACC 2191  
1561 QY TGTGACAGCGGGGACTTACAGCTCAGCCTGCGCGGACGCCGGAAGAACTCTTCAAGAG 1620  
2192 Db TGTGACAGCGGGGACTTACAGCTCAGCCTGCGCGGACGCCGGAAGAACTCTTCAAGAG 2251  
1621 QY AAGTACAGGCCAGCTATGTCGCGAGTCCCTCATCGCTCAGTGGCCCATCAGGTGGAC 1680  
2252 Db AAGTACAGGCCAGCTATGTCGCGAGTCCCTCATCGCTCAGTGGCCCATCAGGTGGAC 2311  
1681 QY GGCAGGGTGTACCACTAGGCTGGTGTGATGCGGCCAGCCCGGAAACCTCAACAGCGG 1740  
2312 Db GGCAGGGTGTACCACTAGGCTGGTGTGATGCGGCCAGCCCGGAAACCTCAACAGCGG 2371  
1741 QY CACTGGCCAGGGGCCCTCAGGACCAAGATGACAAAGATGCTGGGAGCTTTCAGTGGCACT 1800  
2372 Db CACTGGCCAGGGGCCCTCAGGACCAAGATGACAAAGATGCTGGGAGCTTTCAGTGGCACT 2431  
1801 QY GAGAGGCTTCCGACTATCTCAGCGCGCAACCCCATTAAGTGCACATCGGTGTACATC 1860  
2432 Db GAGAGGCTTCCGACTATCTCAGCGCGCAACCCCATTAAGTGCACATCGGTGTACATC 2491  
1861 QY CTAGAGAAACGACACAGTGTGACCTGAGCTGCTGTCAAGTCCCTGAGCGCTGGAAA 1920  
2492 Db CTAGAGAAACGACACAGTGTGACCTGAGCTGCTGTCAAGTCCCTGAGCGCTGGAAA 2551  
1921 QY GACCAACAGCTGCATGACCAACGAGATGAAAACCTTCAGAAACAAATTAAGAACCTG 1980  
2552 Db GACCAACAGCTGCATGACCAACGAGATGAAAACCTTCAGAAACAAATTAAGAACCTG 2611  
1981 QY AGGAGAGTCCGAGGTCACTGAGAGAAAGCGGCCAGAGAGATGTGACGTGCACAAATC 2040  
2612 Db AGGAGAGTCCGAGGTCACTGAGAGAAAGCGGCCAGAGAGATGTGACGTGCACAAATC 2671  
2041 QY AGCTACCAACCCAGCACAAAGCGCCCTCAAGACACAGAGGCTCCAGTCTGCATCTTTT 2100  
2672 Db AGCTACCAACCCAGCACAAAGCGCCCTCAAGACACAGAGGCTCCAGTCTGCATCTTTT 2731  
2101 QY AGGAAGGCTCAGACAGAGAGAGAGAGAGTGTGGCTGTTCGGGAGCAGAGCGCAAGAG 2160  
2732 Db AGGAAGGCTCAGACAGAGAGAGAGAGAGTGTGGCTGTTCGGGAGCAGAGCGCAAGAG 2791  
2161 QY AAACCTCCGCAAGCTGCTCAAGCGCTGACAGAAACACACACAGCATGCGAGGCTC 2220  
2792 Db AAACCTCCGCAAGCTGCTCAAGCGCTGACAGAAACACACACAGCATGCGAGGCTC 2851  
2221 QY AGTGTCTTCAACCAACGACAAACGAGCTGGCAGACGCGCCCTTTCTGGACACTGGGGCT 2280  
2852 Db AGTGTCTTCAACCAACGACAAACGAGCTGGCAGACGCGCCCTTTCTGGACACTGGGGCT 2911  
2281 QY TTCTGTGCTGACACAGCGCCCAACATTAACATGCTGCTGATGAGGACCAATCAATGAG 2340  
2912 Db TTCTGTGCTGACACAGCGCCCAACATTAACATGCTGCTGATGAGGACCAATCAATGAG 2971  
2341 QY ACTCACAATTTCTCTTCTGTGAATTTGCAATGCGGCTCTTAGAGTACTTTGATCTCAAC 2400  
2972 Db ACTCACAATTTCTCTTCTGTGAATTTGCAATGCGGCTCTTAGAGTACTTTGATCTCAAC 3031  
2401 QY ACAGACCCCTACAGCTGATGATGAGTGCAGACACTGACAGGAGTCTCTCAACAG 2460  
3032 Db ACAGACCCCTACAGCTGATGATGAGTGCAGACACTGACAGGAGTCTCTCAACAG 3091  
2461 QY CTACACGTCAGCTCATGGAGCTGAGAGCTGCAAGGGTTTACAAGCAGTGTATACCCCGG 2520



Db	3092	CTACACGTACAGCTCATGAGCTGAGAGCTGCAAGGCTTACAGCAGTGTAAACCCCGG	3151
QY	2521	ACTGGAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGCAGTTT	2580
Db	3152	ACTCGAAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGCAGTTT	3211
QY	2581	CAGCGTGGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGTCGGGACAACTGTGG	2640
Db	3212	CAGCGTGGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGTCGGGACAACTGTGG	3271
QY	2641	GAAGCTGGGAAGGTTAAGAAACAAACAGAGTGAACCTCCAAACAAATGAGGATCACC	2700
Db	3272	GAAGCTGGGAAGGTTAAGAAACAAACAGAGTGAACCTCCAAACAAATGAGGATCACC	3331
QY	2701	TGACTGCACAGGCAATGAAACCAATGTGGTGATTTCCAGAGACCTGTGCTATTGGCC	2760
Db	3332	TGACTGCACAGGCAATGAAACCAATGTGGTGATTTCCAGAGACCTGTGCTATTGGCC	3391
QY	2761	AGGAGGCTGAGAAGCAAGCAACGACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC	2820
Db	3392	AGGAGGCTGAGAAGCAAGCAACGACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC	3451
QY	2821	CAGCAGGACGAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTCTTGGATTA	2880
Db	3452	CAGCAGGACGAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTCTTGGATTA	3511
QY	2881	TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTACCCCTCCC	2940
Db	3512	TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACCACTACCCCTCCC	3571
QY	2941	CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT	3000
Db	3572	CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT	3631
QY	3001	CCCAGGCGGAAAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTCTTTCTAAA	3060
Db	3632	CCCAGGCGGAAAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTCTTTCTAAA	3691
QY	3061	TCCTCTTATTTCTTTGGTTGTCACAAAGAGAACTTAAGAGCAGACAGAGCAACGT	3120
Db	3692	TCCTCTTATTTCTTTGGTTGTCACAAAGAGAACTTAAGAGCAGACAGAGCAACGT	3751
QY	3121	GGAGAGGCTGAAACAGGTGCAGAGAGCTTTCACAAATGAGTCAGTAGCACAAAGAGATGA	3180
Db	3752	GGAGAGGCTGAAACAGGTGCAGAGAGCTTTCACAAATGAGTCAGTAGCACAAAGAGATGA	3811
QY	3181	CATTACCTAGCACTATAAACCTCTGTTGCTCTGAAGAACTCCCTTCATTGTATAT	3240
Db	3812	CATTACCTAGCA-TATAAACCTCTGTTGCTCTGAAGAACTCCCTTCATTGTATAT	3870
QY	3241	GTGACTATTACATGATTAACATGGGAACTTTTAGGGNAACCTATAGAAATCCCAA	3300
Db	3871	GTGACTATTACATGATTAACATGGGAACTTTTAGGGNAACCTATAGAAATCCCAA	3930
QY	3301	TTTTTCAGGAGTGGTGTCTCAATAAACCTCTGTGGCCAGTGTAAAAAGAAAAA	3360
Db	3931	TTTTTCAGGAGTGGTGTCTCAATAAACCTCTGTGGCCAGTGTAAAAAGAAAAA	3990
QY	3361	AA 3362	
Db	3991	AA 3992	
RESULT 5			
AX211371			
LOCUS	AX211371	3348 bp	DNA linear PAT 06-SEP-2001
DEFINITION	Sequence 17 from Patent WO0159127.		
ACCESSION	AX211371		
VERSION	AX211371.1	GI:15523740	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
AUTHORS	Tang, Y.T., Yue, H., Baughn, M.R., Yao, M.G., Bandman, O., Azimzai, V., Lal, P., Gandhi, A.R., Ring, H.Z., Shih, L.L., Yang, J. and Pollock, J.L.		
TITLE	Drug metabolizing enzymes		
JOURNAL	Patent: WO 0159127-A 17 16-AUG-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
Source	Location/Qualifiers		
	1..3348	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/note="Incyte ID No: 1558210CB1"	
BASE COUNT	927 a	933 c	869 g 619 t
ORIGIN			
Query Match	99.5%; Score 3344.8; DB 6; Length 3348;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 3346; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	12	CCCCAAAGAACCCAGATCAGCAAAAAAGAGATGGGCCCCCGAGCCCTCGTCTGTG	71
Db	1	CCCCAAAGAACCCAGATCAGCAAAAAAGAGATGGGCCCCCGAGCCCTCGTCTGTG	60
QY	72	CTTGCTGTCGGCAACTGTGTTCTCCCTGCTGGTGGAAAGCTCGGCTTCTTGTCGACCA	131
Db	61	CTTGCTGTCGGCAACTGTGTTCTCCCTGCTGGTGGAAAGCTCGGCTTCTTGTCGACCA	120
QY	132	CGGCTTGAAGGACAGTTTCAGAGGACCGCAGGAAACATCGCCGCCAACATCATCTGTG	191
Db	121	CGGCTTGAAGGACAGTTTCAGAGGACCGCAGGAAACATCGCCGCCAACATCATCTGTG	180
QY	192	GCTACGGAACGACAGGATGTGGAGTCTGGTTCATGACAGTGATGAACAAGCCGGCG	251
Db	181	GCTACGGAACGACAGGATGTGGAGTCTGGTTCATGACAGTGATGAACAAGCCGGCG	240
QY	252	CATCATGAGACGAGGCGGCGGCGCACTTCATCAAGCCCTTCGTGACCAACCATGTGTG	311
Db	241	CATTATGAGACGAGGCGGCGGCGCACTTCATCAAGCCCTTCGTGACCAACCATGTGTG	300
QY	312	CCCTCAGCTCTCCATCTCCTCAGTGGCAAGTAGTCCAAACCAACAAACCACTACACAA	371
Db	301	CCCTCAGCTCTCCATCTCCTCAGTGGCAAGTAGTCCAAACCAACCAACCACTACACAA	360
QY	372	CAATGAGAACTGCTCTCGCCCTCTGGGACGACGACGAGAGCCGACCTTTGGCGT	431
Db	361	CAATGAGAACTGCTCTCGCCCTCTGGGACGACGACGAGAGCCGACCTTTGGCGT	420
QY	432	GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGATACAA	491
Db	421	GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGATACAA	480
QY	492	CGGCTCTACGTGCCACCGGCTGGAAGAGTGGGTGCGACTCTCTTAAATACTCCGCTT	551
Db	481	CGGCTCTACGTGCCACCGGCTGGAAGAGTGGGTGCGACTCTCTTAAATACTCCGCTT	540
QY	552	TTATAACTACAGCTGTGTGGAAAGGCTGAAAGAAAGCAACGCTCCGACTACTCCAA	611
Db	541	TTATAACTACAGCTGTGTGGAAAGGCTGAAAGAAAGCAACGCTCCGACTACTCCAA	600
QY	612	GGATTACCTACAGACCTCATCAATGACAGCGTGAGCTTTCTTCGCGAGTCCAGAA	671
Db	601	GGATTACCTACAGACCTCATCAATGACAGCGTGAGCTTTCTTCGCGAGTCCAGAA	660
QY	672	GATGTACCGGACAGGCGAGTCTCATGTGTCATGAGCCATGAGCCGCCCGGCGCTGA	731
Db	661	GATGTACCGGACAGGCGAGTCTCATGTGTCATGAGCCATGAGCCGCCCGGCGCTGA	720
QY	732	GGATTACGCCCCCAATATTCACGCTCTTCCCAACGATCTCAGCAATCACGCCGAG	791
Db	721	GGATTACGCCCCCAATATTCACGCTCTTCCCAACGATCTCAGCAATCACGCCGAG	780
QY	792	CTACAACTACGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGNA	851

Db 781 CTACAACTACGGGCCCAACCCGGAACAAACCTTGGATCATGGCTACACGGGGCCCATGAA 840  
QY 852 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCTCATGTC 911  
Db 841 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCTCATGTC 900  
QY 912 GGTGGACGACTCCATGAGACGATTTACCAACATGCTGTTGAGACGGCGGAGCTGGACAA 971  
Db 901 GGTGGACGACTCCATGAGACGATTTACCAACATGCTGTTGAGACGGCGGAGCTGGACAA 960  
QY 972 CAGGTACATCGGTATACACCGCGACACAGGTTTACCACTCGGCCAGTTTGGCCTGGTGA 1031  
Db 961 CAGGTACATCGGTATACACCGCGACACAGGTTTACCACTCGGCCAGTTTGGCCTGGTGA 1020  
QY 1032 AGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGGTTCTACGTGAGGGGCCCAA 1091  
Db 1021 AGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGGTTCTACGTGAGGGGCCCAA 1080  
QY 1092 CGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1151  
Db 1081 CGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1140  
QY 1152 CTTGGACATTTGAGGCTGGACATACCTCGGATATGGACGGGAATCCATCTCAAGCT 1211  
Db 1141 CTTGGACATTTGAGGCTGGACATACCTCGGATATGGACGGGAATCCAATCTCAAGCT 1200  
QY 1212 GCTGGACAGGAGCGGCTGGTGAATCGGTTTCACTTGAAGAAAGATGAGGGTCTGGCG 1271  
Db 1201 GCTGGACAGGAGCGGCTGGTGAATCGGTTTCACTTGAAGAAAGATGAGGGTCTGGCG 1260  
QY 1272 GGAATCTCTTCTGGTGGAGAGGCAAGCTGTACACAGAGAGACAATGACAGAGTGA 1331  
Db 1261 GGAATCTCTTCTGGTGGAGAGGCAAGCTGTACACAGAGAGACAATGACAGAGTGA 1320  
QY 1332 CGCCCGAGGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTGTGTACGCTGC 1391  
Db 1321 CGCCCGAGGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTGTGTACGCTGC 1380  
QY 1392 TGAGTACCAAGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGGAGAGCGCCAC 1451  
Db 1381 TGAGTACCAAGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGGAGAGCGCCAC 1440  
QY 1452 GGGGAAGCTGAAGCTGCATAAGTGAAGGCGCCCATGCGGCTGGGCGGAGCAGAGCCCT 1511  
Db 1441 GGGGAAGCTGAAGCTGCATAAGTGAAGGCGCCCATGCGGCTGGGCGGAGCAGAGCCCT 1500  
QY 1512 CTCGAACCTCGTGCCMAGTACTAGCGGCGAGGCGGAGGCGCTGCACTGTGACAGCGG 1571  
Db 1501 CTCGAACCTCGTGCCMAGTACTAGCGGCGAGGCGGAGGCGCTGCACTGTGACAGCGG 1560  
QY 1572 GGAATCAAGCTCAGCGCTGGCGGAGCGCGGAAACCTCTTCAAGAAAGATGACAGGC 1631  
Db 1561 GGAATCAAGCTCAGCGCTGGCGGAGCGCGGAAACCTCTTCAAGAAAGATGACAGGC 1620  
QY 1632 CAGCTATGTCGGAGTGCCTCCATCGCTCAGTGGCCATCGAGGTGGACGGAGGGTGA 1691  
Db 1621 CAGCTATGTCGGAGTGCCTCCATCGCTCAGTGGCCATCGAGGTGGACGGAGGGTGA 1680  
QY 1692 CCAGTGGCTCGGGTATCGCCCGAGCGCCGAAACCTCACCAAGCGCATCTGGCCAGG 1751  
Db 1681 CCAGTGGCTCGGGTATCGCCCGAGCGCCGAAACCTCACCAAGCGCATCTGGCCAGG 1740  
QY 1752 GGCCCTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACTTGGAGGCTTCC 1811  
Db 1741 GGCCCTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACTTGGAGGCTTCC 1800  
QY 1812 CGACTACTAGCGCCCAACCCCATTAAGTGAACATCGGTGTACATCGGTGTAGAGACGA 1871  
Db 1801 CGACTACTAGCGCCCAACCCCATTAAGTGAACATCGGTGTACATCGGTGTAGAGACGA 1860  
QY 1872 CACAGTCCAGTGTGACCTGACCTGTACAGTCCCTGACGGCTGGAAAGACCAAGCT 1931

Db 1861 CACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGACGGCTGGAAAGACCAAGCT 1920  
QY 1932 GCACATCCGACCAAGATTTGAAACCTCTGACAGAACAAATTAAGAACTTGGAGAACTCG 1991  
Db 1921 GCACATCCGACCAAGATTTGAAACCTCTGACAGAACAAATTAAGAACTTGGAGAACTCG 1980  
QY 1992 AGTCTACCTTGAGAAAGCGGCGAGAGAAATGTGACTGTCAAAATCAGCTACACAC 2051  
Db 1981 AGTCTACCTTGAGAAAGCGGCGAGAGAAATGTGACTGTCAAAATCAGCTACACAC 2040  
QY 2052 CCAGCAAAAGGCGCTCAAGCAGACAGGCTCCAGTCTGCAATCTCTTTCAGGAAGGCGCT 2111  
Db 2041 CCAGCAAAAGGCGCTCAAGCAGACAGGCTCCAGTCTGCAATCTCTTTCAGGAAGGCGCT 2100  
QY 2112 GCAGAGAAAGCAAGGTGTGGCTGTGCGGAGAGAGAGCGCAAGAACTCCGCAA 2171  
Db 2101 GCAGAGAAAGCAAGGTGTGGCTGTGCGGAGAGAGAGCGCAAGAACTCCGCAA 2160  
QY 2172 GCTGCTCAAGCGCTCAGAAACCAACACACAGTGCAGCTGCCAGGCTCCAGCTGCTTAC 2231  
Db 2161 GCTGCTCAAGCGCTCAGAAACCAACACACAGTGCAGCTGCCAGGCTCCAGCTGCTTAC 2220  
QY 2232 CCAGCAAAACAGCACTGGCAGACGGCGCTTTCTTGGACACTGGGCGCTTTCTGTGCTG 2291  
Db 2221 CCAGCAAAACAGCACTGGCAGACGGCGCTTTCTTGGACACTGGGCGCTTTCTGTGCTG 2280  
QY 2292 CACGAGCGCCAAACATTAACACAGTACTGTGTCATGAGGACCATCAATGAGACTCACAA 2351  
Db 2281 CACGAGCGCCAAACATTAACACAGTACTGTGTCATGAGGACCATCAATGAGACTCACAA 2340  
QY 2352 CTTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTA 2411  
Db 2341 CTTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTA 2400  
QY 2412 CCAGCTGATGAATGCAAGTGAACACACTGGACAGGATGTCTTCAACACAGCTACACCTA 2471  
Db 2401 CCAGCTGATGAATGCAAGTGAACACACTGGACAGGATGTCTTCAACACAGCTACACCTA 2460  
QY 2472 GCTCATGGAGCTGAGAGCTGCAAGGTTTACAAGCAGTGTAAACCCCGGACTCGAAACAT 2531  
Db 2461 GCTCATGGAGCTGAGAGCTGCAAGGTTTACAAGCAGTGTAAACCCCGGACTCGAAACAT 2520  
QY 2532 GGAAGCTGGGACTTAAAGATGGAGAAAGCTATGAGCAATACAGGAGTTTCAGGCTGAAA 2591  
Db 2521 GGAAGCTGGGACTTAAAGATGGAGAAAGCTATGAGCAATACAGGAGTTTCAGGCTGAAA 2580  
QY 2592 GTGGCCAGAAATGAAGAGACCTTTCTTCAATCACTGGGCAACTGTGGGAAGGCTGGGA 2651  
Db 2581 GTGGCCAGAAATGAAGAGACCTTTCTTCAATCACTGGGCAACTGTGGGAAGGCTGGGA 2640  
QY 2652 AGGTTAAGAAACCAACAGAGGTGGACCTCCAAAACATAGAGGATCACTGTACTGCAAG 2711  
Db 2641 AGGTTAAGAAACCAACAGAGGTGGACCTCCAAAACATAGAGGATCACTGTACTGCAAG 2700  
QY 2712 GCATGAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTTGGCCAGGAGGCTGA 2771  
Db 2701 GCATGAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTTGGCCAGGAGGCTGA 2760  
QY 2772 GAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAACCCAGCAGAGCA 2831  
Db 2761 GAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAACCCAGCAGAGCA 2820  
QY 2832 GAGATACTTCAGGAAGTCCATTTTGGCTGCTTTGCTTTGGATATATACCTCACCAG 2891  
Db 2821 GAGATACTTCAGGAAGTCCATTTTGGCTGCTTTGCTTTGGATATATACCTCACCAG 2880  
QY 2892 CTGCACAAAATGCATTTTCTGATCAAAAAGTCAACACTTAACCTCCCCAGAGAGCTCA 2951  
Db 2881 CTGCACAAAATGCATTTTCTGATCAAAAAGTCAACACTTAACCTCCCCAGAGAGCTCA 2940  
QY 2952 CAAAGGAAACCGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCTCCCAAGGCGGA 3011  
Db 2941 CAAAGGAAACCGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCTCCCAAGGCGGA 3000

QY 3012 AAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTTAAATCTCTTATTC 3071  
Db |||||  
QY 3001 AAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTTAAATCTCTTATTC 3060  
Db |||||  
QY 3072 TTTTGGTTGTGCACAAAGAGAACTAAGAGCAGGACAGAGGCAACGTGGAGAGGCTGA 3131  
Db |||||  
QY 3061 TTTTGGTTGTGCACAAAGAGAACTAAGAGCAGGACAGAGGCAACGTGGAGAGGCTGA 3120  
Db |||||  
QY 3132 AAACAGTCAGAGAGCGTTTGAACAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAG 3191  
Db |||||  
QY 3121 AAACAGTCAGAGAGCGTTTGAACAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAG 3180  
Db |||||  
QY 3192 CACTATAAACCTGTGTCCTCTGAAGAAACTGCTTCAATGTATATATGTGACTATTTA 3251  
Db |||||  
QY 3181 CACTATAAACCTGTGTCCTCTGAAGAAACTGCTTCAATGTATATATGTGACTATTTA 3240  
Db |||||  
QY 3252 CATGTAATCAACATCGGAACTTTTAGGGGAACTTAATAAGAAATCCCAATTTTTCAGGAGT 3311  
Db |||||  
QY 3241 CATGTAATCAACATCGGAACTTTTAGGGGAACTTAATAAGAAATCCCAATTTTTCAGGAGT 3300  
Db |||||  
QY 3312 GGTGTGTCAATAAACGCTGTGTGCGCCAGTGTAAAGAAATTTTAAAAA 3359  
Db |||||  
QY 3301 GGTGTGTCAATAAACGCTGTGTGCGCCAGTGTAAAGAAATTTTAAAAA 3348  
Db |||||  
RESULT 6  
LOCUS AR252433 3906 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 83 from patent US 6478825.  
ACCESSION AR252433  
VERSION AR252433.1 GI:27300341  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3906)  
AUTHORS Winterbottom J.M., Shimp, L., Boyce, T.M. and Xaes, D.  
TITLE Implant, method of making same and use of the implant for the  
treatment of bone defects  
JOURNAL Patent: US 6478825-A 83 12-NOV-2002;  
FEATURES  
Location/Qualifiers  
source 1..3906  
/organism="unknown"  
BASE COUNT 1047 a 1089 c 1041 g 729 t  
ORIGIN  
Query Match 98.8%; Score 3323.4; DB 6; Length 3906;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;  
QY 1 GACTCCCGCATCCCAAAAGAGCAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60  
Db |||||  
QY 563 GACTCCCGCATCCCAAAAGAGCAGATCAGCAAAAAGAGATGGGCCCCCGAGC 622  
Db |||||  
QY 61 CTCGTCGTGTGCTGTCGCGAACTGTGTCTCCCTGCTGGGTGGAAGCTCGGCTTC 120  
Db |||||  
QY 623 CTCGTCGTGTGCTGTCGCGAACTGTGTCTCCCTGCTGGGTGGAAGCTCGGCTTC 682  
Db |||||  
QY 121 CTGTGCGACCAACCGCTTAAAGGAGAGTTTTCAGAGGGACCGCAGACATCCGCCCCAAC 180  
Db |||||  
QY 693 CTGTGCGACCAACCGCTTAAAGGAGAGTTTTCAGAGGGACCGCAGACATCCGCCCCAAC 742  
Db |||||  
QY 181 ATCATCTCTGGTGTGACGAGCAGACAGGATGTGGAGTGGGTTCCATGCAAGTGATGAAC 240  
Db |||||  
QY 743 ATCATCTCTGGTGTGACGAGCAGACAGGATGTGGAGTGGGTTCCATGCAAGTGATGAAC 802  
Db |||||  
QY 241 AAGACCGGCGCATCATGAGCAGGCGGGGCGGCACTTCATCAACGCTTCGTGACCA 300  
Db |||||  
QY 803 AAGACCGGCGCATCATGAGCAGGCGGGGCGGCACTTCATCAACGCTTCGTGACCA 862  
Db |||||  
QY 301 CCCATGTCGTGCTGCTCAGCTCTCCATCTCTCACTGCGCAAGTACGTCCCAACCAAC 360  
Db |||||

Db 863 CCCATGTCGTGCTGCTCAGCTCTCCATCTCTCACTGCGCAAGTACGTCCCAACCAACAAAC 922  
QY |||||  
QY 361 ACCTACACCAACAAAGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGCCGC 420  
Db |||||  
QY 923 ACCTACACCAACAAAGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGCCGC 982  
Db |||||  
QY 421 ACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480  
Db |||||  
QY 983 ACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 1042  
Db |||||  
QY 481 AATGATACAAAGGCTCTTACCTGCGACCCCGCTGGAAGAGTGGTGGACTCTCTTAA 540  
Db |||||  
QY 1043 AATGATACAAAGGCTCTTACCTGCGACCCCGCTGGAAGAGTGGTGGACTCTCTTAA 1102  
Db |||||  
QY 541 AACTCCCGCTTTTATAAATACACGCTGTGTGCGAAGCGGGTGAAGAAAGCAACGCTCC 600  
Db |||||  
QY 1103 AACTCCCGCTTTTATAAATACACGCTGTGTGCGAAGCGGGTGAAGAAAGCAACGCTCC 1162  
Db |||||  
QY 601 GACTACTCCAGGATTAACCTCAACAGCTCATCAACATGACAGCTGAGCTTCTTCGCG 660  
Db |||||  
QY 1163 GACTACTCCAGGATTAACCTCAACAGCTCATCAACATGACAGCTGAGCTTCTTCGCG 1222  
Db |||||  
QY 661 AGCTCCAGAGATGTACCCGACAGGCGAGCTCTCATGCTCATGAGCAATCAGCCGCC 720  
Db |||||  
QY 1223 AGCTCCAGAGATGTACCCGACAGGCGAGCTCTCATGCTCATGAGCAATCAGCCGCC 1282  
Db |||||  
QY 721 CACGCGCTGAGGATTCAGCCCAACATTAATCAAGCTCTTTCCCAACGCAATCTCAGCAC 780  
Db |||||  
QY 1283 CACGCGCTGAGGATTCAGCCCAACATTAATCAAGCTCTTTCCCAACGCAATCTCAGCAC 1342  
Db |||||  
QY 781 ATCAGCGCGAGCTCAACATGAGCGGCCCAACCGGCAACACTGATCATGCGCTTACAGG 840  
Db |||||  
QY 1343 ATCAGCGCGAGCTCAACATGAGCGGCCCAACCGGCAACACTGATCATGCGCTTACAGG 1402  
Db |||||  
QY 841 GGGCCCATGAAGCCCATCCACATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCAG 900  
Db |||||  
QY 1403 GGGCCCATGAAGCCCATCCACATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCAG 1462  
Db |||||  
QY 901 ACCCTCATGTGCGTGGAGCTCCATGAGAGAGATTTACAACTGCTGTTGAGAGCGGC 960  
Db |||||  
QY 1463 ACCCTCATGTGCGTGGAGCTCCATGAGAGAGATTTACAACTGCTGTTGAGAGCGGC 1522  
Db |||||  
QY 961 GAGCTGGCAACACGCTACATCGTATACACCGCGGACCAACGCTTACCATCGGCCAGTTT 1020  
Db |||||  
QY 1523 GAGCTGGCAACACGCTACATCGTATACACCGCGGACCAACGCTTACCATCGGCCAGTTT 1582  
Db |||||  
QY 1021 GGCCTGTGAAAGGAAATTCATGCTCATGATGATGATGATGATGATGATGATGATGATG 1080  
Db |||||  
QY 1583 GGCCTGTGAAAGGAAATTCATGCTCATGATGATGATGATGATGATGATGATGATGATG 1642  
Db |||||  
QY 1081 AGGGGCCCAACGCTGGAAGCGGCTGCTGAATCCCACTGCTCTCAACATGACCTG 1140  
Db |||||  
QY 1643 AGGGGCCCAACGCTGGAAGCGGCTGCTGAATCCCACTGCTCTCAACATGACCTG 1702  
Db |||||  
QY 1141 GCGCCCAACATCCTCGACATTCAGGCTGAGCATACTCGCGGATATGAGCGGAAATTC 1200  
Db |||||  
QY 1703 GCGCCCAACATCCTCGACATTCAGGCTGAGCATACTCGCGGATATGAGCGGAAATTC 1762  
Db |||||  
QY 1201 ATCTCTAAGCTGCTGGAACCGGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1260  
Db |||||  
QY 1763 ATCTCTAAGCTGCTGGAACCGGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1822  
Db |||||  
QY 1261 AGGGCTCTGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACAAAGAGAGCAAT 1320  
Db |||||  
QY 1823 AGGGCTCTGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACAAAGAGAGCAAT 1882  
Db |||||  
QY 1321 GACAGGTGAGAGCGCCCGAGAGAGAGATTTCTGCGCAAGTACCAAGCTGTGAGAGACCTG 1380  
Db |||||  
QY 1883 GACAGGTGAGAGCGCCCGAGAGAGAGATTTCTGCGCAAGTACCAAGCTGTGAGAGACCTG 1942  
Db |||||  
QY 1381 TGTGAGCGTGTGAGTACAGAGCGGCTGTGAGAGCTGCGGACAGAGAGTGGCAGTGTGTG 1440  
Db |||||  
QY 1943 TGTGAGCGTGTGAGTACAGAGCGGCTGTGAGAGCTGCGGACAGAGAGTGGCAGTGTGTG 2002  
Db |||||

QY	1441	GAGGAGCGCCACGGGGAGCTGAAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGC	1500
Db	2003	GAGGAGCGCCACGGGGAGCTGAAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGC	2062
QY	1501	AGCAGAGGCTCTTCACACCTTCGTCGCCCAAGTACTACGGGCAGAGGCGACGAGGGCCTGCACC	1560
Db	2063	AGCAGAGGCTCTTCACACCTTCGTCGCCCAAGTACTACGGGCAGAGGCGACGAGGGCCTGCACC	2122
QY	1561	TGTGACAGCGGGGACATCAAGAGCTCAGCTGGCGGGAGCGCGGAAAAAACTCTTTCAAGAAG	1620
Db	2123	TGTGACAGCGGGGACATCAAGAGCTCAGCTGGCGGGAGCGCGGAAAAAACTCTTTCAAGAAG	2182
QY	1621	AAGTACAAAGSCCAGCATATGTCGCGAGTCGCTCCATCCGCTCAGTGGGCCATCAGAGGTGGAC	1680
Db	2183	AAGTACAAAGSCCAGCATATGTCGCGAGTCGCTCCATCCGCTCAGTGGGCCATCAGAGGTGGAC	2242
QY	1681	GGCAGGGTGTAACAAGTAGGCTGGGTGATGCGCGCCACGCGCCGAAACCTTCACCAAGCGG	1740
Db	2243	GGCAGGGTGTAACAAGTAGGCTGGGTGATGCGCGCCACGCGCCGAAACCTTCACCAAGCGG	2302
QY	1741	CAGTCGCCAGGGGCCCTGTAGGACCAAGATGACAAAGATGGTGGGAGCTTTCAGTGGCACT	1800
Db	2303	CAGTCGCCAGGGGCCCTGTAGGACCAAGATGACAAAGATGGTGGGAGCTTTCAGTGGCACT	2362
QY	1801	GGAGGGCTTCCCAGTACTCAGCGCGCCAAACCCCATTAAGTGTACATCGGTGCTACATC	1860
Db	2363	GGAGGGCTTCCCAGTACTCAGCGCGCCAAACCCCATTAAGTGTACATCGGTGCTACATC	2422
QY	1861	CTAGAGAACACACAGTCAGTGTGACCTGGAGCTGTACAGGTCCCTGCAGGCGCTCGAAA	1920
Db	2423	CTAGAGAACACACAGTCAGTGTGACCTGGAGCTGTACAGGTCCCTGCAGGCGCTCGAAA	2482
QY	1921	GACCACAAGCTGCACATCGACACGAGATGTAACCCCTGCAGAACAAAAATTAAGAACCTTG	1980
Db	2483	GACCACAAGCTGCACATCGACACGAGATGTAACCCCTGCAGAACAAAAATTAAGAACCTTG	2542
QY	1981	AGGGAAGTCGAGGTCACTGAGAAAAAGCGCGCAGAGAAATGTGACTGTCACAAAATC	2040
Db	2543	AGGGAAGTCGAGGTCACTGAGAAAAAGCGCGCAGAGAAATGTGACTGTCACAAAATC	2602
QY	2041	AGGTACCAACCCAGCAGAAAGGCGCCTCAAGCAGAGAGGCTCCAGTCTGCATCCTTTTC	2100
Db	2603	AGGTACCAACCCAGCAGAAAGGCGCCTCAAGCAGAGAGGCTCCAGTCTGCATCCTTTTC	2662
QY	2101	AGGAGGGCTGCAAGAGAGGACAGGTGTGGCTGTTCGGGAGCAGAGCCGCAAGAG	2160
Db	2663	AGGAGGGCTGCAAGAGAGGACAGGTGTGGCTGTTCGGGAGCAGAGCCGCAAGAG	2722
QY	2161	AAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACAAAGACAGTGCAGCATGCCAGGCGCTC	2220
Db	2723	AAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACAAAGACAGTGCAGCATGCCAGGCGCTC	2782
QY	2221	ACGTGCTTCAACCAAGCAACCAAGCATGGCAGACGCGCGCTTTCTGGACATCTGGGGCCT	2280
Db	2783	ACGTGCTTCAACCAAGCAACCAAGCATGGCAGACGCGCGCTTTCTGGACATCTGGGGCCT	2842
QY	2281	TTCTGTGCCTGCAACCGCGCAACATTAACACGTACTGTGTGCATGAGGACCATCAATGAG	2340
Db	2843	TTCTGTGCCTGCAACCGCGCAACATTAACACGTACTGTGTGCATGAGGACCATCAATGAG	2902
QY	2341	ACTCACAATTTCTCTGTGTGAATTTGCAATTTGGCTTCTTAGAGTACTTTGATCTCAAC	2400
Db	2903	ACTCACAATTTCTCTGTGTGAATTTGCAATTTGGCTTCTTAGAGTACTTTGATCTCAAC	2962
QY	2401	ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGACACAGGAGTGTCTTCAACAG	2460
Db	2963	ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGACACAGGAGTGTCTTCAACAG	3022
QY	2461	CTACACGTAAGCTCATGGAGCTGAGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGG	2520
Db	3023	CTACACGTAAGCTCATGGAGCTGAGAGCTGCAAGGGTTTCAAGCAGTGTAAACCCCGG	3082

Qy	2521	ACTCGA	AACATG	AGCCTG	GGGACCTT	AAAGAT	TGGAGG	AAAGCT	ATGAGCA	ATACAGG	CAGTTT	2588
Db	3083	ACTCGA	AACATG	AGCCTG	GGGACCTG	-----	GATGGAG	GAAGCT	TATGAGCA	ATACAGG	CAGTTT	3133
Qy	2581	CAGCGT	CGAAAG	TGGCCG	AGAAAT	TGAAGAG	CACCTT	CTTCCAA	ATCT	TGGGACA	CAACTGTG	2640
Db	3134	CAGCGT	CGAAAG	TGGCCG	AGAAAT	TGAAGAG	CACCTT	CTTCCAA	ATCT	TGGGACA	CAACTGTG	3193
Qy	2641	GAAGGT	TGGGAAG	GGTTAA	GAAACAA	CAGAGG	TGGGAC	CTTCCAAA	AAACAT	TAGAGGC	ATCAC	2700
Db	3194	GAAGGT	TGGGAAG	GGTTAA	GAAACAA	CAGAGG	TGGGAC	CTTCCAAA	AAACAT	TAGAGGC	ATCAC	3253
Qy	2701	TGACTC	GCACAG	GCACAA	TGA AAAA	CCATGT	CGGTGAT	TTCAG	CAGAG	CACCTG	CTATTGG	2760
Db	3254	TGACTC	GCACAG	GCACAA	TGA AAAA	CCATGT	CGGTGAT	TTCAG	CAGAG	CACCTG	CTATTGG	3313
Qy	2761	AGGAGC	CTTGAG	AAACCA	AGCAGC	GCACCT	CTCAGT	CAACAT	GTACAG	ATTCG	GAGGATA	2820
Db	3314	AGGAGC	CTTGAG	AAACCA	AGCAGC	GCACCT	CTCAGT	CAACAT	GTACAG	ATTCG	GAGGATA	3373
Qy	2821	CAGCAG	GCAGAG	TAACTT	CAGGAG	TGCATTTT	TGCCCTG	CTTTTGC	TTTGCCTT	TGCTTTG	CGTTTGA	2880
Db	3374	CAGCAG	GCAGAG	TAACTT	CAGGAG	TGCATTTT	TGCCCTG	CTTTTGC	TTTGCCTT	TGCTTTG	CGTTTGA	3433
Qy	2881	TACCTC	ACACAG	CTGCAC	AAATGCA	TTTTTGGT	ATCAAAA	AGTCA	CCACTA	CAACCTC	CCCTCC	2940
Db	3434	TACCTC	ACACAG	CTGCAC	AAATGCA	TTTTTGGT	ATCAAAA	AGTCA	CCACTA	CAACCTC	CCCTCC	3493
Qy	2941	CCAGAA	AGCTCA	CAAGG	AAAA	CGAGAG	AGCGAG	CGAGAG	AGATTT	CCCTTGG	AAATTTCT	3000
Db	3494	CCAGAA	AGCTCA	CAAGG	AAAA	CGAGAG	AGCGAG	CGAGAG	AGATTT	CCCTTGG	AAATTTCT	3553
Qy	3001	CCCAAG	GGCGAA	AGTCAT	TGGGAATTTT	TAAATCAT	ATAGG	GGGAAA	AGCAGT	CTGTCT	TCTAAA	3060
Db	3554	CCCAAG	GGCGAA	AGTCAT	TGGGAATTTT	TAAATCAT	ATAGG	GGGAAA	AGCAGT	CTGTCT	TCTAAA	3613
Qy	3061	TCCTCT	TATTTCT	TTTGGT	TTGTG	CAAGA	AGGA	ACTA	AGAG	CGAGGA	CAGAGCG	3120
Db	3614	TCCTCT	TATTTCT	TTTGGT	TTGTG	CAAGA	AGGA	ACTA	AGAG	CGAGGA	CAGAGCG	3673
Qy	3121	GGAGAG	CGTGA	AAACAG	TGCAGAG	AGCTTT	TGACAT	GTAGT	CAGTAG	CACAAA	AGAGATGA	3180
Db	3674	GGAGAG	CGTGA	AAACAG	TGCAGAG	AGCTTT	TGACAT	GTAGT	CAGTAG	CACAAA	AGAGATGA	3733
Qy	3181	CATTAC	CTAGCA	CTATAA	ACCCCTG	GTGCTCTG	TAAG	AAACCTG	CTTCA	TTGTATAT	TAT	3240
Db	3734	CATTAC	CTAGCA	CTATAA	ACCCCTG	GTGCTCTG	TAAG	AAACCTG	CTTCA	TTGTATAT	TAT	3793
Qy	3241	GTGACT	ATTATAC	ATGTAAT	CAACAT	GGGAACTTTT	TAGG	GGGAACCTT	ATAG	AAATCC	CA	3300
Db	3794	GTGACT	ATTATAC	ATGTAAT	CAACAT	GGGAACTTTT	TAGG	GGGAACCTT	ATAG	AAATCC	CA	3853
Qy	3301	TTTTT	CAGGAG	TGGTGGTGT	CAATAA	ACCGCTCTG	TGGC	CAGTGT	TAAAG	AAAA	3353	
Db	3854	TTTTT	CAGGAG	TGGTGGTGT	CAATAA	ACCGCTCTG	TGGC	CAGTGT	TAAAG	AAAA	3906	

RESULT 7

AX403196

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Ashkenazi, A.J., Baker, K.P., Botstein, D., Desnovers, J., Eaton, D., Ferrara, N., Gerber, H., Gerritsen, M., Goddard, A., Godowski, P., Grimaldi, C.J., Guney, A.L., Kljavin, I., Napiet, M.A., Pan, J., Paoni, N.F., Roy, M., Stewart, T.A., Tumas, D., Watanabe, C.K.,

AX403196.1

GI:21436787

Sequence 83 from Patent WO0073454.

3906 bp

DNA

linear

PAT 14-JUN-2002

Williams, P., Wood, W.I. and Zhang, Z.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
the same  
Patent: WO 0073454-A 83 07-DEC-2000;  
Genentech Inc. (US)

FEATURES  
source location/Qualifiers  
1..3906  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 1047 a 1089 c 1041 g 729 t  
ORIGIN

Query Match 98.8%; Score 3323.4; DB 6; Length 3906;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

1 GACTCCCGCATCCCAAGAGACACAGATCAGCAAAAAGAGATGGGCCCGGAGC 60  
563 GACTCCCGCATCCCAAGAGACACAGATCAGCAAAAAGAGATGGGCCCGGAGC 622  
61 CTGCTGCTGCTTCTGCTCGCAACTGTGTCTCTCTGCTGGGTGGAAGCTCGGCTTC 120  
623 CTGCTGCTGCTTCTGCTCGCAACTGTGTCTCTCTGCTGGGTGGAAGCTCGGCTTC 682  
121 CTGTCGACACCGCTGAAAGGACAGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 180  
683 CTGTCGACACCGCTGAAAGGACAGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 742  
181 ATCATCTCTGCTGCTGAGGACGACAGAGATGTGGAGCTGGGTTCATGCAAGTGATGAAC 240  
743 ATCATCTCTGCTGCTGAGGACGACAGAGATGTGGAGCTGGGTTCATGCAAGTGATGAAC 802  
241 AAGACCGCGGCATCATGAGACGAGGCGGCGCGCATTCATCAACGGCTTCGTGACCA 300  
803 AAGACCGCGGCATCATGAGACGAGGCGGCGCGCATTCATCAACGGCTTCGTGACCA 862  
301 CCCATGTGCTGCCCTCAGCTCCTCATCTCTCACTGCGCAAGTACGTCACACCAAC 360  
863 CCCATGTGCTGCCCTCAGCTCCTCATCTCTCACTGCGCAAGTACGTCACACCAAC 922  
361 ACCTACACCAACATGAACTGTCTCTGCGCTCTCTGGCAGGACACAGCAGAGCGCG 420  
923 ACCTACACCAACATGAACTGTCTCTGCGCTCTCTGGCAGGACACAGCAGAGCGCG 982  
421 ACCTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTCTTCGGAGATCTT 480  
983 ACCTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTCTTCGGAGATCTT 1042  
481 AATGAATACACCGCTCTACGTGCCACCGGCTGGAGAGAGTGGGTGCGACTCCTTAA 540  
1043 AATGAATACACCGCTCTACGTGCCACCGGCTGGAGAGAGTGGGTGCGACTCCTTAA 1102  
541 AACTCCCGCTTTTAACTACACGCTGTGCGAAACCGGGTGAAGAAAGCAGCGCTCC 600  
1103 AACTCCCGCTTTTAACTACACGCTGTGCGAAACCGGGTGAAGAAAGCAGCGCTCC 1162  
601 GACTACTCCAGAGATTAACCTCACAGACTCATACCAATGACACGCTGAGCTTCTCCG 660  
1163 GACTACTCCAGAGATTAACCTCACAGACTCATACCAATGACACGCTGAGCTTCTCCG 1222  
661 ACCTCCAGAGAGATGATACCCGACAGCGCAAGTCTCATGTGTCATGACCCATGAGCCCC 720  
1223 ACCTCCAGAGAGATGATACCCGACAGCGCAAGTCTCATGTGTCATGACCCATGAGCCCC 1282  
721 CACGCGCTGAGGATTCAGCCGCAATATTCAGCTCTTCCCAACGCTCTCAGCAC 780  
1283 CACGCGCTGAGGATTCAGCCGCAATATTCAGCTCTTCCCAACGCTCTCAGCAC 1342  
781 ATCAGCGGAGCTACAACTAGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACAG 840  
1343 ATCAGCGGAGCTACAACTAGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACAG 1402

QY 841 GGGCCCATGAAGCCCATCCCATGAAATTCACCAATGCTCCAGCGAAGCGTGTGAG 900  
DB 1403 GGGCCCATGAAGCCCATCCCATGAAATTCACCAATGCTCCAGCGAAGCGTGTGAG 1462  
QY 901 ACCCTCATGTGGTGGAGGACTCCATGGAGAGAGATTTACACATGCTGTTGAGACGGC 960  
DB 1463 ACCCTCATGTGGTGGAGGACTCCATGGAGAGAGATTTACACATGCTGTTGAGACGGC 1522  
QY 961 GAGCTGGAGCAACAGGTATCATCGTATACACCGCCGACCAACCGTTCACCATCGGCAGTTT 1020  
DB 1523 GAGCTGGAGCAACAGGTATCATCGTATACACCGCCGACCAACCGTTCACCATCGGCAGTTT 1582  
QY 1021 GGCCTGTGTAAGGGAATTCATGCGCATATGAGTTTGACATCAGGGTCCCTTCTACGTG 1080  
DB 1583 GGCCTGTGTAAGGGAATTCATGCGCATATGAGTTTGACATCAGGGTCCCTTCTACGTG 1642  
QY 1081 AGGGCCCCCAACAGGTGGAAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTG 1140  
DB 1643 AGGGCCCCCAACAGGTGGAAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTG 1702  
QY 1141 GGCCTCCACATCTTGGACATTTGCAAGGCTGTGACATATCTGCGGATATGGACGGGAATTC 1200  
DB 1703 GGCCTCCACATCTTGGACATTTGCAAGGCTGTGACATATCTGCGGATATGGACGGGAATTC 1762  
QY 1201 ATCTCTAAGCTGTGACACAGGCGGCGGTGATCGGTTTCACTTGAAAAAGAGATG 1260  
DB 1763 ATCTCTAAGCTGTGACACAGGCGGCGGTGATCGGTTTCACTTGAAAAAGAGATG 1822  
QY 1261 AGGGTCTTGGCGGACTCTCTTCTTGTGGAGAGAGCGCAAGCTGTACACAGAGAGACAAT 1320  
DB 1823 AGGGTCTTGGCGGACTCTCTTCTTGTGGAGAGAGCGCAAGCTGTACACAGAGAGACAAT 1882  
QY 1321 GACAAAGTGGAGCGCCAGGAGAGAACTTTCTGCGCAAGTACCAAGCTGTGTAAGGACCTG 1380  
DB 1883 GACAAAGTGGAGCGCCAGGAGAGAACTTTCTGCGCAAGTACCAAGCTGTGTAAGGACCTG 1942  
QY 1381 TGTGACGCTGTGAGTACAGACGCGGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440  
DB 1943 TGTGACGCTGTGAGTACAGACGCGGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2002  
QY 1441 GAGGACGCGACGGGGAAGCTGAAGCTGAAGTGAAGGCGCCCATGCGGCTGGGCGGC 1500  
DB 2003 GAGGACGCGACGGGGAAGCTGAAGCTGAAGTGAAGGCGCCCATGCGGCTGGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCCAACTGTCGCCCAAGTACTACCGGCGAGGCGAGCGCTTGACCC 1560  
DB 2063 AGCAGAGCCCTCTCCAACTGTCGCCCAAGTACTACCGGCGAGGCGAGCGCTTGACCC 2122  
QY 1561 TGTGACGCGGGAAGTCAAGCTCAGCTGCGCGGACGCGGAAAAAACTCTTCAAGAG 1620  
DB 2123 TGTGACGCGGGAAGTCAAGCTCAGCTGCGCGGACGCGGAAAAAACTCTTCAAGAG 2182  
QY 1621 AAGTACAGGCGCAGCTATGTCGCGAGTGCCTGCAATCGCTCAGTGGCCATCGAGTGGAC 1680  
DB 2183 AAGTACAGGCGCAGCTATGTCGCGAGTGCCTGCAATCGCTCAGTGGCCATCGAGTGGAC 2242  
QY 1681 GGCAGGCTGTACCACTAGGCTGCGGTGATCGCGCCGAGCCCGAAAACTCACCAGCGG 1740  
DB 2243 GGCAGGCTGTACCACTAGGCTGCGGTGATCGCGCCGAGCCCGAAAACTCACCAGCGG 2302  
QY 1741 CACTGGCCAGGGGCGCTGTGAGGACCAAGATGAAGGATGTGGGAGCTTCAGTGGGACT 1800  
DB 2303 CACTGGCCAGGGGCGCTGTGAGGACCAAGATGAAGGATGTGGGAGCTTCAGTGGGACT 2362  
QY 1801 GGAGGCTTTCGCGACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 1860  
DB 2363 GGAGGCTTTCGCGACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 2422  
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGAAGCTGTACAGTCCCTGAGGCTTGAAA 1920  
DB 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGAAGCTGTACAGTCCCTGAGGCTTGAAA 2482  
QY 1921 GACCACAGCTGCATCGACCAAGGATTTGAAACCTTGACAGAACAAATTAAGAACCTG 1980









Db 2903 ACTCACAATTTCTCTTCTGTGAATTTTGCAACTGGCTTCTAGAGTACTTTGTATCTCAAC 2962  
QY 2401 ACAGACCCCTACAGCTGATGAATGTCAGTGAACACACTGGACAGGATGCTCTCAACACAG 2460  
Db 2963 ACAGACCCCTACAGCTGATGAATGTCAGTGAACACACTGGACAGGATGCTCTCAACAG 3022  
QY 2461 CTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAAGCAGTGTAAACCCCGG 2520  
Db 3023 CTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAAGCAGTGTAAACCCCGG 3082  
QY 2521 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAGGAGCTGATGAGCAATACAGCAGTTT 2580  
Db 3083 ACTCGAAACATGGACCTG-----GATGAGGAGGAGCTGATGAGCAATACAGCAGTTT 3133  
QY 2581 CAGCGTCGAAGTGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACAACTGTGG 2640  
Db 3134 CAGCGTCGAAGTGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACAACTGTGG 3193  
QY 2641 GAAGGCTGGGAGGTTAAGAAACACAGAGGTGAGCTCCAAACATAGAGGATCAAC 2700  
Db 3194 GAAGGCTGGGAGGTTAAGAAACACAGAGGTGAGCTCCAAACATAGAGGATCAAC 3253  
QY 2701 TGACTGCAGGCAATGAAACACCATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760  
Db 3254 TGACTGCAGGCAATGAAACACCATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313  
QY 2761 AGGAGGCTGGAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATAAC 2820  
Db 3314 AGGAGGCTGGAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATAAC 3373  
QY 2821 CAGCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3433  
QY 2881 TACCTCAACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACATAACCTCTCC 2940  
Db 3434 TACCTCAACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACATAACCTCTCC 3493  
QY 2941 CCAGAGGCTCAACAGGAAACCGGAGAGCAGCAGGAGAGATTTCTTGGAAATTTCT 3000  
Db 3494 CCAGAGGCTCAACAGGAAACCGGAGAGCAGCAGGAGAGATTTCTTGGAAATTTCT 3553  
QY 3001 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTTAAA 3060  
Db 3554 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTTAAA 3613  
QY 3061 TCCTCTTATCTTTTGGTTGTCAAAAAGGAACTTAAGAGCAGCAGCAGAGGCAAGT 3120  
Db 3614 TCCTCTTATCTTTTGGTTGTCAAAAAGGAACTTAAGAGCAGCAGCAGAGGCAAGT 3673  
QY 3121 GGAGAGGCTGAAAACAGTGCAGAGACGTTTGACATGAGTCAAGTACGACAAAAGAGATGA 3180  
Db 3674 GGAGAGGCTGAAAACAGTGCAGAGACGTTTGACATGAGTCAAGTACGACAAAAGAGATGA 3733  
QY 3181 CATTTACCTAGCATTATAAACCCTGGTTCCTCTGAAGAACTCGCTTCATGTATATAT 3240  
Db 3734 CATTTACCTAGCATTATAAACCCTGGTTCCTCTGAAGAACTCGCTTCATGTATATAT 3793  
QY 3241 GTGACTATTATACATGTAATCAACATGGGAACTTTTGGGGAACCTTAATAGAAATCCCAA 3300  
Db 3794 GTGACTATTATACATGTAATCAACATGGGAACTTTTGGGGAACCTTAATAGAAATCCCAA 3853  
QY 3301 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3353  
Db 3854 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3906

RESULT 9  
AX463886  
LOCUS AX463886  
DEFINITION Sequence 19 from Patent WO0140466.  
ACCESSION AX463886  
VERSION AX463886.1 GI:21898960

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E., Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L., Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K., Wood, W.L. and Zhang, Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0140466-A 19 JUN-2001; Genentech Inc. (US)  
FEATURES  
Location/Qualifiers  
Source 1..3906  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 1047 a 1089 c 1041 g 729 t  
ORIGIN  
Query Match 98.8%; Score 3323.4; DB 6; Length 3906;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3343; Conservative 1; Mismatches 9; Gaps 1;  
QY 1 GACTCCCGCATCCCAAAAGAGCACCAGATCAGCAAAAAGAGAGTGGCCCCCGAGC 60  
Db 563 GACTCCCGCATCCCAAAAGAGCACCAGATCAGCAAAAAGAGAGTGGCCCCCGAGC 622  
QY 61 CTGCTGCTGTGCTGCTGCTGCGCAACTGTGTTCTCCCTGCTGGGTGGAGCTCGCCCTTC 120  
Db 623 CTGCTGCTGTGCTGCTGCTGCGCAACTGTGTTCTCCCTGCTGGGTGGAGCTCGCCCTTC 682  
QY 121 CTGCTGCAACCGCTCGAAAGCAGGTTTCAGAGGACCGCAGGAAACATCGCCCCAAC 180  
Db 683 CTGCTGCAACCGCTCGAAAGCAGGTTTCAGAGGACCGCAGGAAACATCGCCCCAAC 742  
QY 181 ATCATCTCTGCTCTCAGCAGCAGCAGCAGATGTGGAGCTGGGTTCATCGAGTGATGAAC 240  
Db 743 ATCATCTCTGCTCTCAGCAGCAGCAGCAGATGTGGAGCTGGGTTCATCGAGTGATGAAC 802  
QY 241 AAGACCCCGCGCATCATGAGAGCAGGCGGGGCGCACTTCATCAACGCTTCGTGACCA 300  
Db 803 AAGACCCCGCGCATCATGAGAGCAGGCGGGGCGCACTTCATCAACGCTTCGTGACCA 862  
QY 301 CCCATGTGCTGCCCTCAGCTCCCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
Db 863 CCCATGTGCTGCCCTCAGCTCCCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 922  
QY 361 ACCTACACCAACAATGAGAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
Db 923 ACCTACACCAACAATGAGAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982  
QY 421 ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGAGATATCT 480  
Db 983 ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGAGATATCT 1042  
QY 481 AATGAATACAAACGGCTCTTACGTCGCAACCCGCTCGAAGAGAGTGGGTCTCGGACTCT 540  
Db 1043 AATGAATACAAACGGCTCTTACGTCGCAACCCGCTCGAAGAGAGTGGGTCTCGGACTCT 1102  
QY 541 ACTCCCGCTTTTATTAATATACAGCTGTGTGGAACGGGGTGAAGAAAGCAGCGCTCC 600  
Db 1103 ACTCCCGCTTTTATTAATATACAGCTGTGTGGAACGGGGTGAAGAAAGCAGCGCTCC 1162  
QY 601 GACTACTTCAAGGATTTACCTCAGACACTCATCAACATGACAGCTGAGCTTCTTCGCG 660  
Db 1163 GACTACTTCAAGGATTTACCTCAGACACTCATCAACATGACAGCTGAGCTTCTTCGCG 1222  
QY 661 ACCTCAAGAAAGATGTACCCGACAGGCGAGTCTCTCATGCTCATCAGCCATCAGCCGCC 720  
Db 1223 ACCTCAAGAAAGATGTACCCGACAGGCGAGTCTCTCATGCTCATCAGCCATCAGCCGCC 1282

QY 721 CACGGCCCTGAGGATTCAGCCCCCAATATTACGGCTCTTCCCAACGGCATCTCAGCAC 780  
Db 1283 CACGGCCCTGAGGATTCAGCCCCCAATATTACGGCTCTTCCCAACGGCATCTCAGCAC 1342  
QY 781 ATCAGCGCGAGCTCAACTAGCGCCGCCAACCGGACAAACACTGGATCATCGCGTACAG 840  
Db 1343 ATCAGCGCGAGCTCAACTAGCGCCGCCAACCGGACAAACACTGGATCATCGCGTACAG 1402  
QY 841 GGGCCCATGAAGCCCATCCATCATGGAATTTACCAACATGCTCCAGCGGAAGCGCTTCAG 900  
Db 1403 GGGCCCATGAAGCCCATCCATCATGGAATTTACCAACATGCTCCAGCGGAAGCGCTTCAG 1462  
QY 901 ACCCTCATGTGCGTGGAGCACTCATGAGAGCAATTTACAACTGCTGGTTGAGACGGGC 960  
Db 1463 ACCCTCATGTGCGTGGAGCACTCATGAGAGCAATTTACAACTGCTGGTTGAGACGGGC 1522  
QY 961 GAGCTGGACAACACATCATCGTATACACCGCCGACACGGTTACCAATCGCGCAGTTT 1020  
Db 1523 GAGCTGGACAACACATCATCGTATACACCGCCGACACGGTTACCAATCGCGCAGTTT 1582  
QY 1021 GGCTGTGTGAAGGAAATCATGCGCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1080  
Db 1583 GGCTGTGTGAAGGAAATCATGCGCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1642  
QY 1081 AGGGGCCCCACGTGGAGCGGCTGTGTAATCCCGACATCGTCTCAACATGACCTG 1140  
Db 1643 AGGGGCCCCACGTGGAGCGGCTGTGTAATCCCGACATCGTCTCAACATGACCTG 1702  
QY 1141 GCCCCACCATCTGTGACATTTGACGGCTTGACATCTCGGATATGAGCGGGAATTC 1200  
Db 1703 GCCCCACCATCTGTGACATTTGACGGCTTGACATCTCGGATATGAGCGGGAATTC 1762  
QY 1201 ATCTCTAAGCTGTGACACGGAGCGCGGTGTAATCGGTTTCACTTGAAGAAAGAGATG 1260  
Db 1763 ATCTCTAAGCTGTGACACGGAGCGCGGTGTAATCGGTTTCACTTGAAGAAAGAGATG 1822  
QY 1261 AGGCTCTGGCGGACTCCTCTTGGTGGAGAGCGCAAGTGTGTACACAGAGAGACAAT 1320  
Db 1823 AGGCTCTGGCGGACTCCTCTTGGTGGAGAGCGCAAGTGTGTACACAGAGAGACAAT 1882  
QY 1321 GACAAGGTGGACGCGCAGGAGGAACTTTCTGCCAAGTACCAAGCGTGTGAAGGACCTG 1380  
Db 1883 GACAAGGTGGACGCGCAGGAGGAACTTTCTGCCAAGTACCAAGCGTGTGAAGGACCTG 1942  
QY 1381 TGTGAGCGTGTGAGTACCAAGCGCGTGTGAGCAGCTGGGACGAGTGGCAGTGTG 1440  
Db 1943 TGTGAGCGTGTGAGTACCAAGCGCGTGTGAGCAGCTGGGACGAGTGGCAGTGTG 2002  
QY 1441 GAGGACGCGCAGCGGGAAGCTGAAGCTGCAATAGTGCAAGGGCCCCCATGCGGCTGGGCGGC 1500  
Db 2003 GAGGACGCGCAGCGGGAAGCTGAAGCTGCAATAGTGCAAGGGCCCCCATGCGGCTGGGCGGC 2062  
QY 1501 AGCAGAGCCCTCTCAACCTGTGCCCAAGTACTACGGGACGGGACGAGCGCTGACAC 1560  
Db 2063 AGCAGAGCCCTCTCAACCTGTGCCCAAGTACTACGGGACGGGACGAGCGCTGACAC 2122  
QY 1561 TGTGACAGCGGGACTACAAGCTCAGCGCTGCGCGGACGCGGAAAAAATCTTTCAAGAAG 1620  
Db 2123 TGTGACAGCGGGACTACAAGCTCAGCGCTGCGCGGACGCGGAAAAAATCTTTCAAGAAG 2182  
QY 1621 AAGTACAAGGCGCAGTATGTCCGAGTGTGCTTCATCCGCTCAGTGGCCATCGAGGTGGAC 1680  
Db 2183 AAGTACAAGGCGCAGTATGTCCGAGTGTGCTTCATCCGCTCAGTGGCCATCGAGGTGGAC 2242  
QY 1681 GGCAGGGTGTACACGTAGGCGCTGGGTGATGTCGCCGCCAGCCCGAATCTCACCAAGCGG 1740  
Db 2243 GGCAGGGTGTACACGTAGGCGCTGGGTGATGTCGCCGCCAGCCCGAATCTCACCAAGCGG 2302  
QY 1741 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAAGGATGGTGGGACTTTCAGTGGCACT 1800  
Db 2303 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAAGGATGGTGGGACTTTCAGTGGCACT 2362

QY 1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCAATATAAGTGACACATCGTGCTACATC 1860  
Db 2363 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCAATATAAGTGACACATCGTGCTACATC 2422  
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAAGTCCCTGAGGCTTGAAAA 1920  
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAAGTCCCTGAGGCTTGAAAA 2482  
QY 1921 GACCAACAGCTGCACATCGACCACGAGATTGAAACCTCTGCAGAACAAAAATTAAAGAACCTG 1980  
Db 2483 GACCAACAGCTGCACATCGACCACGAGATTGAAACCTCTGCAGAACAAAAATTAAAGAACCTG 2542  
QY 1981 AGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCGCAGAGAATGTGATGTCACAAAAATC 2040  
Db 2543 AGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCGCAGAGAATGTGATGTCACAAAAATC 2602  
QY 2041 AGCTACCAACCCACGACAAAGGGCGCTCAAGGACACAGAGCTCCAGTCTGCTCTTC 2100  
Db 2603 AGCTACCAACCCACGACAAAGGGCGCTCAAGGACACAGAGCTCCAGTCTGCTCTTC 2662  
QY 2101 AGGAAGGGCTGCAAGAGAAGGACAAAGGTGTGGCTGTTCGGGAGCAGAAGCGCAAGAG 2160  
Db 2663 AGGAAGGGCTGCAAGAGAAGGACAAAGGTGTGGCTGTTCGGGAGCAGAAGCGCAAGAG 2722  
QY 2161 AAATCCCGAGCTGCTCAAGCGCTGCGACACACGACACGTCGATGCCAGGCTC 2220  
Db 2723 AAATCCCGAGCTGCTCAAGCGCTGCGACACACGACACGTCGATGCCAGGCTC 2782  
QY 2221 AGTGCTTTCACCCACGACAAACAGCACTGGGAGAGCGGCGCTTCTGGACACTGGGGCT 2280  
Db 2783 AGTGCTTTCACCCACGACAAACAGCACTGGGAGAGCGGCGCTTCTGGACACTGGGGCT 2842  
QY 2281 TTCTGTGCTGACACGAGCGCAACAAATAAACGTAATCTGGTGCATGAGGACCAATCAATGAG 2340  
Db 2843 TTCTGTGCTGACACGAGCGCAACAAATAAACGTAATCTGGTGCATGAGGACCAATCAATGAG 2902  
QY 2341 ACTCAAAATTTCTCTCTGTGAAATTTGCAACTGCTTCCCTAGAGTACTTTGATCTCAAC 2400  
Db 2903 ACTCAAAATTTCTCTCTGTGAAATTTGCAACTGCTTCCCTAGAGTACTTTGATCTCAAC 2962  
QY 2401 ACAGACCCCTACAGCTGATGAATGACAGTGAACACACTGGAACGGGATGTCTCAACCCAG 2460  
Db 2963 ACAGACCCCTACAGCTGATGAATGACAGTGAACACACTGGAACGGGATGTCTCAACCCAG 3022  
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGCTGCAAGGTTACAGAGTGTAAACCCCGG 2520  
Db 3023 CTACAGCTACAGCTCATGAGCTGAGGCTGCAAGGTTACAGAGTGTAAACCCCGG 3082  
QY 2521 ACTCGAAACATGGACCTTGAAGATGGAGGAAGCTATGAGCAATACAGGCACTTT 2580  
Db 3083 ACTCGAAACATGGACCTG-----GATGGAGGAAGCTATGAGCAATACAGGCACTTT 3133  
QY 2581 CAGCGTCAAGAGTGGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640  
Db 3134 CAGCGTCAAGAGTGGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 3193  
QY 2641 GAAGGCTGGGAAGGTTAAGAAACCAACAGAGGTGACCTCCAAAAACATAGAGGCACTACC 2700  
Db 3194 GAAGGCTGGGAAGGTTAAGAAACCAACAGAGGTGACCTCCAAAAACATAGAGGCACTACC 3253  
QY 2701 TGACTGCAAGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760  
Db 3254 TGACTGCAAGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313  
QY 2761 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 2820  
Db 3314 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 3373  
QY 2821 CAGCAGGAGCAGAGATTAATCTTCAGAAAGTCCATTTTGGCCCTGCTTTTGTGTTGATTA 2880  
Db 3374 CAGCAGGAGCAGAGATTAATCTTCAGAAAGTCCATTTTGGCCCTGCTTTTGTGTTGATTA 3433  
QY 2881 TACCTCACCGCTGCACAAAAATGATTTTTTCGTATCAAAAAAGTACCACCTAACCCCTCC 2940













..CATGTGCTGCCCCCTCAGCGCTCTCCATCTCCTCAGCGCAAGTACGTCCCAACCAAC 991  
121 ACCTACCAACCAATGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGAGCGCG 420  
992 ACCTACCAACCAATGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGAGCGCG 1051  
421 ACCTTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTTCTTCGGAGAGTATCTT 480  
1052 ACCTTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTTCTTCGGAGAGTATCTT 1111  
481 AATGATACAAACGGCTCTTACGTGCAACCGGCTGGAAGAGTGGGTGCACTCTTAA 540  
1112 AATGATACAAACGGCTCTTACGTGCAACCGGCTGGAAGAGTGGGTGCACTCTTAA 1171  
541 AACTCCCGCTTTTATPACTACACGCTGTGTGGAACGGGTGGAAGAAAGACAGCGCTCC 600  
1172 AACTCCCGCTTTTATPACTACACGCTGTGTGGAACGGGTGGAAGAAAGACAGCGCTCC 1231  
601 GACTACTCCAAAGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 660  
1232 GACTACTCCAAAGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 1291  
661 ACCTCCAAAGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 720  
1292 ACCTCCAAAGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 1351  
721 CAGCGCCCTGAGGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 780  
1352 CAGCGCCCTGAGGATTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 1411  
781 ATCAACCGGAGCTTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 840  
1412 ATCAACCGGAGCTTAACCTCAGACAGCTCATACCAATGACAGGTGAGCTTTCCGC 1471  
841 GGGCCCATGAGGCCATTCACATGGAATTCACCAATGCTTCCAGCGGAGAGCGCTTTCAG 900  
1472 GGGCCCATGAGGCCATTCACATGGAATTCACCAATGCTTCCAGCGGAGAGCGCTTTCAG 1531  
901 ACCTCATGTGCGTGAAGCTTCCATGAGAGAGATTTACACATGCTGCTGAGAGCGCG 960  
1532 ACCTCATGTGCGTGAAGCTTCCATGAGAGAGATTTACACATGCTGCTGAGAGCGCG 1591  
961 GAGCTGGAACAACAGTACATGTAACCGCGGACCAACCGTTACACATCGGCCAGTTT 1020  
1592 GAGCTGGAACAACAGTACATGTAACCGCGGACCAACCGTTACACATCGGCCAGTTT 1651  
1021 GGCCTGTGGAAGGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1652 GGCCTGTGGAAGGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1711  
1081 AGGGGCCCCAAGTGAAGCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1140  
1712 AGGGGCCCCAAGTGAAGCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1771  
1141 GCGCCCAACCATCTGACATTTGAGGCTGGAACATACCTGCGGATATGGAACGGGAATCC 1200  
1772 GCGCCCAACCATCTGACATTTGAGGCTGGAACATACCTGCGGATATGGAACGGGAATCC 1831  
1201 ATCTCAAGCTGCTGACAGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 1260  
1832 ATCTCAAGCTGCTGACAGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 1891  
1261 AGGGTGTGGGGAGCTCTTCTTGTGTGAGAGAGGCAAGCTGCTTACCAAGAGACAAT 1320  
1892 AGGGTGTGGGGAGCTCTTCTTGTGTGAGAGAGGCAAGCTGCTTACCAAGAGACAAT 1951  
1321 GACAAGGTGGAACCGCCAGGAGGAGAACTTTCTGCGCAAGTACAGAGTGTGGAAGACCTG 1380  
1952 GACAAGGTGGAACCGCCAGGAGGAGAACTTTCTGCGCAAGTACAGAGTGTGGAAGACCTG 2011  
1381 TGTACAGCTGCTGAGTACAGAGCGCTGTGAGAGCTGGAAGAGTGTGAGAGTGTGAG 1440  
2012 TGTACAGCTGCTGAGTACAGAGCGCTGTGAGAGCTGGAAGAGTGTGAGAGTGTGAG 2071

QY 1441 GAGGAGCCCAACGGGGAAGCTGAAGCTGATAGTGAAGGGCCCCCATGCGGCTGCGCGC 1500  
DB 2072 GAGGAGCCCAACGGGGAAGCTGAAGCTGATAGTGAAGGGCCCCCATGCGGCTGCGCGC 2131  
QY 1501 AGCAGAGCCCTTCCAACTCGTCCCAAGTACTACGGGAGGGGAGGAGGCTGAC 1560  
DB 2132 AGCAGAGCCCTTCCAACTCGTCCCAAGTACTACGGGAGGGGAGGAGGCTGAC 2191  
QY 1561 TGTGACAGCGGGGACTACAACTCAGCTCGGCGGAGCGCGGAAAAAACTCTTCAAGAG 1620  
DB 2192 TGTGACAGCGGGGACTACAACTCAGCTCGGCGGAGCGCGGAAAAAACTCTTCAAGAG 2251  
QY 1621 AAGTACAGAGCGGAGCTATGTCGCGAGCTCGCTCCATCGCTCAGTGGCCATCGAGGTGAG 1680  
DB 2252 AAGTACAGAGCGGAGCTATGTCGCGAGCTCGCTCCATCGCTCAGTGGCCATCGAGGTGAG 2311  
QY 1681 GGCAGGGGTGTACAGTATGAGGCTGGGTGATGCGCGCCAGCGCCGAAAACTCTCAAGCGG 1740  
DB 2312 GGCAGGGGTGTACAGTATGAGGCTGGGTGATGCGCGCCAGCGCCGAAAACTCTCAAGCGG 2371  
QY 1741 CACTGGCCAGAGGGGCTTGAAGGACCAAGATGACAAAGATGACAAAGATGACAAAGATGAC 1800  
DB 2372 CACTGGCCAGAGGGGCTTGAAGGACCAAGATGACAAAGATGACAAAGATGACAAAGATGAC 2431  
QY 1801 GAGGCGCTTCCCGAGCTACTCAGCGCGCAACCCCAATTAAGTGAACATCGGTGCTACATC 1860  
DB 2432 GAGGCGCTTCCCGAGCTACTCAGCGCGCAACCCCAATTAAGTGAACATCGGTGCTACATC 2491  
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAAAGTCCCTGAGCGCTGGAAA 1920  
DB 2492 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAAAGTCCCTGAGCGCTGGAAA 2551  
QY 1921 GACCAAGAGCTGACATCGACACCGAGGATGAAACCTCTGAGAAACAAATTAAGAACCTG 1980  
DB 2552 GACCAAGAGCTGACATCGACACCGAGGATGAAACCTCTGAGAAACAAATTAAGAACCTG 2611  
QY 1981 AGGAGCTCCGAGCTTCACTGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040  
DB 2612 AGGAGCTCCGAGCTTCACTGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2671  
QY 2041 AGCTACCAACACCCAGGCTTCAAGGCGGCTCAAGACAGAGGCTCCAGTCTGATCTTTC 2100  
DB 2672 AGCTACCAACACCCAGGCTTCAAGGCGGCTCAAGACAGAGGCTCCAGTCTGATCTTTC 2731  
QY 2101 AGGAGGCGCTGCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160  
DB 2732 AGGAGGCGCTGCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2791  
QY 2161 AAATCCGCAAGCTGCTCAAGCGGCTGAGAAACACGACAGTGAAGATGCAAGGCTC 2220  
DB 2792 AAATCCGCAAGCTGCTCAAGCGGCTGAGAAACACGACAGTGAAGATGCAAGGCTC 2851  
QY 2221 AGGTGCTTCAACCCAGGCTTCAAGCGGCTGAGAAACACGACAGTGAAGATGCAAGGCTC 2280  
DB 2852 AGGTGCTTCAACCCAGGCTTCAAGCGGCTGAGAAACACGACAGTGAAGATGCAAGGCTC 2911  
QY 2281 TTCTGTGCTGCAACGCGGCAACATTAACAGTACTGCTGATGATGATGATGATGATGATGAT 2340  
DB 2912 TTCTGTGCTGCAACGCGGCAACATTAACAGTACTGCTGATGATGATGATGATGATGATGAT 2971  
QY 2341 ACTCAAAATTTCTTCTGTAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2400  
DB 2972 ACTCAAAATTTCTTCTGTAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 3031  
QY 2401 ACAGAGCCCTTACAGCTGATGAATGCAAGTGAACACACTGAGAGGATGCTCTCAACAG 2460  
DB 3032 ACAGAGCCCTTACAGCTGATGAATGCAAGTGAACACACTGAGAGGATGCTCTCAACAG 3091  
QY 2461 CTACAGTACAGCTCATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2520  
DB 3092 CTACAGTACAGCTCATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 3151





661 AGCCATGACGCCCCCAACGCGCCCTGAGGATTACGCCCAACAATATTACGCGCTCTTCCCA 720  
766 AAGCATCTCAGACATACACGCGAGCTCAACTACGCGCCCAACCCGGAACAACTGG 825  
721 AAGCATCTCAGACATACACGCGAGCTCAACTACGCGCCCAACCCGGAACAACTGG 780  
826 ATCATCGCTTACACGCGGCCCATGAAGCCCATCCCATGGAATTCACCAACATGCTCCAG 885  
781 ATCATCGCTTACACGCGGCCCATGAAGCCCATCCCATGGAATTCACCAACATGCTCCAG 840  
886 CGAAGCGCTTGAGACCTCATGTCGGTGGAGACTCATGAGAGAGAGATTTACAACTG 945  
841 CGAAGCGCTTGAGACCTCATGTCGGTGGAGACTCATGAGAGAGAGATTTACAACTG 900  
946 CTGGTTGAGACGCGCGAGCTGGACACAGTACATCGTATACACGCGCCGACACCGTTAC 1005  
901 CTGGTTGAGACGCGCGAGCTGGACACAGTACATCGTATACACGCGCCGACACCGTTAC 960  
1006 CACATCGCGCAGTTTGGCTGGTGAAGGGAATCCATGCCATATGAGTTTGACATCAGG 1065  
961 CACATCGCGCAGTTTGGCTGGTGAAGGGAATCCATGCCATATGAGTTTGACATCAGG 1020  
1066 GTCCGCTTACCTGAGGCGCCCAAGCTGGAGCGGCTGTCTGAATCCCCACATCGTC 1125  
1021 GTCCGCTTACCTGAGGCGCCCAAGCTGGAGCGGCTGTCTGAATCCCCACATCGTC 1080  
1126 CTCAACATTCACCTGGCGCCCAACATCTGGAATTCAGGCGCTGACATCCTCGGAT 1185  
1081 CTCAACATTCACCTGGCGCCCAACATCTGGAATTCAGGCGCTGACATCCTCGGAT 1140  
1186 ATGGAACGGAATTCATCTCAAGCTGGACACGAGCGCGGCTGTCTGAATCCCCACATCGTC 1245  
1141 ATGGAACGGAATTCATCTCAAGCTGGACACGAGCGCGGCTGTCTGAATCCCCACATCGTC 1200  
1246 TTGAAAAGAGATGAGGCTTGGCGGACTCTCTTGTGGAGAGAGCAAGCTGCTA 1305  
1201 TTGAAAAGAGATGAGGCTTGGCGGACTCTCTTGTGGAGAGAGCAAGCTGCTA 1260  
1306 CACAAGAGAGACAATGACAGGTTGGACGCGCCAGAGAGAACTTTCTGCCAAGTACCAG 1365  
1261 CACAAGAGAGACAATGACAGGTTGGACGCGCCAGAGAGAACTTTCTGCCAAGTACCAG 1320  
1366 CGTGTGAAGACCTGTCTCAGCTGTGAGTACAGACGCGGTGTGAGCAGCTGGACAG 1425  
1321 CGTGTGAAGACCTGTGTACGCTGTGAGTACAGACGCGGTGTGAGCAGCTGGACAG 1380  
1426 AAGTGGCAGTGTGTGAGGACGCGCACGCGGGAAGCTGAAGTGCATAGTGCAGAGGCC 1485  
1381 AAGTGGCAGTGTGTGAGGACGCGCACGCGGGAAGCTGAAGTGCATAGTGCAGAGGCC 1440  
1486 ATGGGCTGGGCGGACGAGCGCTCTCCACTCTGTCGCCAAGTACTACGGCGAGGGC 1545  
1441 ATGGGCTGGGCGGACGAGCGCTCTCCACTCTGTCGCCAAGTACTACGGCGAGGGC 1500  
1546 AGCAGGCGCTCAGCTGTGACAGCGGGAAGTACAGCTCAGCTGGCGGACGCGCGAAA 1605  
1501 AGCAGGCGCTCAGCTGTGACAGCGGGAAGTACAGCTCAGCTGGCGGACGCGCGAAA 1560  
1606 AAATCTTTCAAGAGAGAGTACAGGCGAGCTATGTCGAGTGCCTCCATCCGCTCAGTG 1665  
1561 AAATCTTTCAAGAGAGAGTACAGGCGAGCTATGTCGAGTGCCTCCATCCGCTCAGTG 1620  
1666 GCCATCAGGTTGGACGCGGAGGCTGTACACGCTAGGCTGGGTATGCGGCCAGGCGCGA 1725  
1621 GCCATCAGGTTGGACGCGGAGGCTGTACACGCTAGGCTGGGTATGCGGCCAGGCGCGA 1680  
1726 AACCTCACCAAGCGGACATGCGGAGGCGGCTGTGAGACCAAGATGACAAGATGATGGG 1785  
1681 AACCTCACCAAGCGGACATGCGGAGGCGGCTGTGAGACCAAGATGACAAGATGATGGG 1740  
1786 GACTTCAGTGGACATGAGGCTTCCGACTACTCAGCGGCAACCCATTAAGTACACA 1845  
1741 GACTTCAGTGGACATGAGGCTTCCGACTACTCAGCGGCAACCCATTAAGTACACA 1800

1846 CATCGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTGACCTGTGTAAGTCC 1905  
1801 CATCGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTGACCTGTGTAAGTCC 1860  
1906 CTGAGCGCTTGGAAAGACCAACAGTGTGACATCGACCAAGATTTGAACCTTGCAGAAC 1965  
1861 CTGAGCGCTTGGAAAGACCAACAGTGTGACATCGACCAAGATTTGAACCTTGCAGAAC 1920  
1966 AATAATTAAGAACCTGAGGGAAGTCCGAGGTCACTGAAGAAAAGCGGCCAGAGAATGT 2025  
1921 AATAATTAAGAACCTGAGGGAAGTCCGAGGTCACTGAAGAAAAGCGGCCAGAGAATGT 1980  
2026 GACTGTCACAAAATCAGTACCAACACCCAGCAAAAAGCGGCCCTCAAGCACAGAGGCTCC 2085  
1981 GACTGTCACAAAATCAGTACCAACACCCAGCAAAAAGCGGCCCTCAAGCACAGAGGCTCC 2040  
2086 AGTCTGCATCTTTTCAAGGAGGCTTCAAGAGAGGACAAAGTGTGGCTGTTCGGGAG 2145  
2041 AGTCTGCATCTTTTCAAGGAGGCTTCAAGAGAGGACAAAGTGTGGCTGTTCGGGAG 2100  
2146 CAGAAGCGCAAGAAACTCCGCAAGCTCTCAAGCGCTCGAGAAACAAACGACACGTCG 2205  
2101 CAGAAGCGCAAGAAACTCCGCAAGCTCTCAAGCGCTCGAGAAACAAACGACACGTCG 2160  
2206 AGCATGCGAGGCTTCAAGTGTTCACCCAGCAACCCAGCACTGGCAGAGCGGCGCTTC 2265  
2161 AGCATGCGAGGCTTCAAGTGTTCACCCAGCAACCCAGCACTGGCAGAGCGGCGCTTC 2220  
2266 TGGACACTTGGGCGCTTCTGTGCTGCACGAGCGCCAAACAAATTAACACGTAATGTCATG 2325  
2221 TGGACACTTGGGCGCTTCTGTGCTGCACGAGCGCCAAACAAATTAACACGTAATGTCATG 2280  
2326 AGGACCATCAATAGAGACTCAATTTCTCTTGTGAATTTGCAACTGGCTTCTTAGAG 2385  
2281 AGGACCATCAATAGAGACTCAATTTCTCTTGTGAATTTGCAACTGGCTTCTTAGAG 2340  
2386 TACTTTGATCTCAACACAGACCCCTTACGAGCTGATGAATGCAAGTGAACACACTGCAAGG 2445  
2341 TACTTTGATCTCAACACAGACCCCTTACGAGCTGATGAATGCAAGTGAACACACTGCAAGG 2400  
2446 GATGCTCTCAACACAGCTACAGTACAGTATGAGCTGAGGAGCTGCAAGGTTACAAG 2505  
2401 GATGCTCTCAACACAGCTACAGTACAGTATGAGCTGAGGAGCTGCAAGGTTACAAG 2460  
2506 CAGTGTAAACCCCGGACTCGAAGACATGAGCTGAGCTTAAAGATGAGGAAGCTATGAG 2565  
2461 CAGTGTAAACCCCGGACTCGAAGACATGAGCTGAGCTTAAAGATGAGGAAGCTATGAG 2520  
2566 CAATACAGGCGAGTTTCAAGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCA 2625  
2521 CAATACAGGCGAGTTTCAAGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCA 2580  
2626 CTGGAGCAACTGTGGGAAGGCTGGAAGGT 2655  
2581 CTGGAGCAACTGTGGGAAGGCTGGAAGGT 2610

## RESULT 15

BD127910

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 2559)

Ota.T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,

Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and

Koga,H.

BD127910 2559 bp DNA linear PAT 18-SEP-2002  
Primer for synthesizing full-length cDNA and use thereof.



TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002017375-A 3341 22-JAN-2002;  
HELIIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

PN JP 2002017375-A/3341

PD 22-JAN-2002

PF 07-JUL-2000 JP 2000253172

PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO

PI ISHII,

PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI

SHINICHI KOJIMA,

PI TETSUJI OTSUKI,HISASHI KOGA

PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC

10,

PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT CDS (133)..(1602).

Location/Qualifiers

1..2559

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 732 a 635 c 666 g 526 t

ORIGIN

Query Match 68.2%; Score 2294.8; DB 6; Length 2559;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2296; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1055 TTGACATCAGGTCCTCCGTTCTACGTCAGGCGGCCCAAGTCGGAAGCGCGTCTGTAATC 1114

Db 2 TGGACATCAGGTCCTCCGTTCTACGTCAGGCGGCCCAAGTCGGAAGCGCGTCTGTAATC 61

QY 1115 CCCACATCGTCTTCAACATTCGACCTGGCCCCCAACATCTCGACATTCGAGCCCTGGACA 1174

Db 62 CCCACATCGTCTTCAACATTCGACCTGGCCCCCAACATCTCGACATTCGAGCCCTGGACA 121

QY 1175 TACCTCGGATATGACGGGAATCATCTCAGCTGCTGGACACGGAGCGCGGTGA 1234

Db 122 TACCTCGGATATGACGGGAATCATCTCAGCTGCTGGACACGGAGCGCGGTGA 181

QY 1235 ATCGGTTTCACTTGAAGAGATGAGGGTCTGGCGGACTCTCTTCTTGTGGAGAGAG 1294

Db 182 ATCGGTTTCACTTGAAGAGATGAGGGTCTGGCGGACTCTCTTCTTGTGGAGAGAG 241

QY 1295 GCAAGTCTACACAGAGAGCAATGACAGGTGGACGCCAGAGAGAACTTCTGC 1354

Db 242 GCAAGTCTACACAGAGAGCAATGACAGGTGGACGCCAGAGAGAACTTCTGC 301

QY 1355 CCAAGTACCAAGGCTGTGAAGGACCTGTGTCAAGCGTGTGAGTACCAAGCGCGTGTGAGC 1414

Db 302 CCAAGTACCAAGGCTGTGAAGGACCTGTGTCAAGCGTGTGAGTACCAAGCGCGTGTGAGC 361

QY 1415 AGCTGGGACAGAAGTGGCAGTGTGTGGAGAGCGCCACCGGGAAGCTGAAGTGTGATAAGT 1474

Db 362 AGCTGGGACAGAAGTGGCAGTGTGTGGAGAGCGCCACCGGGAAGCTGAAGTGTGATAAGT 421

QY 1475 GCAAGGCCCCATCGGCTGGGCGGACAGAGCCCTCTCCAACTCGTGCCTCAAGTACT 1534

Db 422 GCAAGGCCCCATCGGCTGGGCGGACAGAGCCCTCTCCAACTCGTGCCTCAAGTACT 481

QY 1535 ACGGCGAGGCGAGCGGCTTCCACTGTGACAGCGGGGACTTACAAGCTCAGCTGGCCG 1594

Db 482 ACGGCGAGGCGAGCGGCTTCCACTGTGACAGCGGGGACTTACAAGCTCAGCTGGCCG 541

QY 1595 GACGCGGGAATCTCTCAAGAGAGATCAAGGCGAGCTATGTCGCGAGTGCCTCCA 1654

Db 542 GACGCGGGAATCTCTCTCAAGAGAGATCAAGGCGAGCTATGTCGCGAGTGCCTCCA 601

QY 1655 TCCGCTCAGTGGCCATCCAGGTGGAGCGGCGAGGGGTGTACCACTAGGCTGGGTGATGCCG 1714

Db 1714

602 TCCGCTCAGTGGCCATCGAGGTGGACGGGGGTGTACCAAGTATGGCTGGTGTGATGCCG 661

QY 1715 CCCAGCCCCGAAACCTCTACCAAGCGGCACTGGCCAGGGGCGCCCTGAGGACCAAGATGACA 1774

Db 662 CCCAGCCCCGAAACCTCTACCAAGCGGCACTGGCCAGGGGCGCCCTGAGGACCAAGATGACA 721

QY 1775 AGGATGGTGGGACCTTCAGTGGCACTGGAGGCTTCCGACATCTCAGCCGCGCAACCCCA 1834

Db 722 AGGATGGTGGGACCTTCAGTGGCACTGGAGGCTTCCGACATCTCAGCCGCGCAACCCCA 781

QY 1835 TTAAGTGGACACATCGGTGCTACATCTTAGAGAACGACACAGTCCAGTGTGACCTTGGACC 1894

Db 782 TTAAGTGGACACATCGGTGCTACATCTTAGAGAACGACACAGTCCAGTGTGACCTTGGACC 841

QY 1895 TGTACAAGTCCCTGCAGGCTGGAAAGACCAAGCTGCACATCGACCAAGATTTGAAA 1954

Db 842 TGTACAAGTCCCTGCAGGCTGGAAAGACCAAGCTGCACATCGACCAAGATTTGAAA 901

QY 1955 CCTTCGAGAACAAATTAAGACCTTGGAGGAGTCCGAGGTCACTGAAGAAAGAGCGC 2014

Db 902 CCTTCGAGAACAAATTAAGACCTTGGAGGAGTCCGAGGTCACTGAAGAAAGAGCGC 961

QY 2015 CAGAAAGATGTGACTGTCTACAAAATCAGCTACCAACCCAGCAGCAAAAGCGCCCTCAAGC 2074

Db 962 CAGAAAGATGTGACTGTCTACAAAATCAGCTACCAACCCAGCAGCAAAAGCGCCCTCAAGC 1021

QY 2075 ACAGAGGCTCAGTCTGCATCTTTTCAGAAAGGGCTGCAAGAGAGCAAGGTGTGGC 2134

Db 1022 ACAGAGGCTCAGTCTGCATCTTTTCAGAAAGGGCTGCAAGAGAGCAAGGTGTGGC 1081

QY 2135 TGTTCGGGAGCAGAAAGCGCAAGAAATCCGCAAGCTGCTCAAGCGCTGCAAGACA 2194

Db 1082 TGTTCGGGAGCAGAAAGCGCAAGAAATCCGCAAGCTGCTCAAGCGCTGCAAGACA 1141

QY 2195 AGCACAAGTGCAGATGCGAGGCTTCAAGTGTCTTCAACCAAGCAACCAAGCACTGGCAGA 2254

Db 1142 AGCACAAGTGCAGATGCGAGGCTTCAAGTGTCTTCAACCAAGCAACCAAGCACTGGCAGA 1201

QY 2255 CGGCGCTTTCTGGACACTTGGGGCTTCTGTGCTGCAAGCGCCCAACCAATACAGT 2314

Db 1202 CGGCGCTTTCTGGACACTTGGGGCTTCTGTGCTGCAAGCGCCCAACCAATACAGT 1261

QY 2315 ACTGCTGATGAGGACCAATCAATGAGACTCAATTTCTTCTTGTGAATTTGCAATG 2374

Db 1262 ACTGCTGATGAGGACCAATCAATGAGACTCAATTTCTTCTTGTGAATTTGCAATG 1321

QY 2375 GCTTCTAGAGTACTTGTATCTCAACAGACCCCTTACAGCTGATGATGAGTCAAGTGAACA 2434

Db 1322 GCTTCTAGAGTACTTGTATCTCAACAGACCCCTTACAGCTGATGATGAGTCAAGTGAACA 1381

QY 2435 CACTGGACAGGATGTCTTCAACAGCTTACAGTCAAGCTCATGGAGCTGAGGAGCTGCA 2494

Db 1382 CACTGGACAGGATGTCTTCAACAGCTTACAGTCAAGCTCATGGAGCTGAGGAGCTGCA 1441

QY 2495 AGGTTTACAGAGTGTAAACCCCGGACCTCGAACAATGGAAGCTGTAAGATGGAG 2554

Db 1442 AGGTTTACAGAGTGTAAACCCCGGACCTCGAACAATGGAAGCTGTAAGATGGAG 1501

QY 2555 GAAGCTATCAGCAATACAGGCAAGTTTCAAGCTGCAAAAGTGGCCAGAAATGAAGAGACCTT 2614

Db 1502 GAAGCTATCAGCAATACAGGCAAGTTTCAAGCTGCAAAAGTGGCCAGAAATGAAGAGACCTT 1561

QY 2615 CTTTCCAAATCCTTGGGACAACTGTGGGAAGGCTGGGAAGGTTTGAAGAAACAAGAGGTTGG 2674

Db 1562 CTTTCCAAATCCTTGGGACAACTGTGGGAAGGCTGGGAAGGTTTGAAGAAACAAGAGGTTGG 1621

QY 2675 ACCTCCAAACATAGAGGCATCACTGACTCAGAGCAATGAAGAAACAATGTGGTGA 2734

Db 1622 ACCTCCAAACATAGAGGCATCACTGACTCAGAGCAATGAAGAAACAATGTGGTGA 1681

QY 2735 TTTCCAGCAGACTGTGCTTATTTGGCCAGAGGCTTGAGAAAGCAAGCACTCTCAGT 2794

Db 1682 TTTCCAGCAGACTGTGCTTATTTGGCCAGAGGCTTGAGAAAGCAAGCACTCTCAGT 1741



QY	2795	CAACATGACAGATTCTGGAGGATAACACGACGAGCAGAGATAA	2854
Db	1742	CAACATGACAGATTCTGGAGGATAACACGACGAGCAGAGATAA	1801
QY	2855	TTTGGCCCTGCTTTTGGATTATACCTCACCAGCTGCACAAATG	2914
Db	1802	TTTGGCCCTGCTTTTGGATTATACCTCACCAGCTGCACAAATG	1861
QY	2915	ATCAAAAGTCAACCACTAACCCCTCCCCAGAGCTCACAAGGAA	2974
Db	1862	ATCAAAAGTCAACCACTAACCCCTCCCCAGAGCTCACAAGGAA	1921
QY	2975	CGAGAGATTTCCTTGGAAATTTCTCCAGGGCGAAAGTCATTG	3034
Db	1922	CGAGAGATTTCCTTGGAAATTTCTCCAGGGCGAAAGTCATTG	1981
QY	3035	TAGGGAAAAGCAGTCTCTTCTTAATCCTCTTATTTCTTTTGT	3094
Db	1982	TAGGGAAAAGCAGTCTCTTCTTAATCCTCTTATTTCTTTTGT	2041
QY	3095	ACTAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAACAGT	3154
Db	2042	ACTAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAACAGT	2101
QY	3155	ATGAGTCAGTAGCACAAAGAGATGACATTTTACCTAGCACTA	3214
Db	2102	ATGAGTCAGTAGCACAAAGAGATGACATTTTACCTAGCACTA	2161
QY	3215	GAAGAACTGCCTTCATTGTATATATGTGACTATTTACATGTA	3274
Db	2162	GAAGAACTGCCTTCATTGTATATATGTGACTATTTACATGTA	2221
QY	3275	TAGGGGAACCTTAATAGAAATCCCAATTTTACGAGTGGTGTCA	3334
Db	2222	TAGGGGAACCTTAATAGAAATCCCAATTTTACGAGTGGTGTCA	2281
QY	3335	GGCCAGTGTAAAGAAAA	3352
Db	2282	GGCCAGTGTAAAGAAAA	2299

Search completed: February 14, 2004, 17:54:31  
Job time : 11919 secs



```
QY 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDQDVELGS 60
DB 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDQDVELGS 60
QY 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTKYVHNHNTYNNENCSFSSWA 120
DB 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTKYVHNHNTYNNENCSFSSWA 120
QY 121 QHESRTFAVINSTGYRTAFPGKYLNEYNGSVYPPGKKEWGLLKNSRFYNTLCRNGVK 180
DB 121 QHESRTFAVINSTGYRTAFPGKYLNEYNGSVYPPGKKEWGLLKNSRFYNTLCRNGVK 180
QY 181 EKHSYSDYLDLTLDLNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
DB 181 EKHSYSDYLDLTLDLNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
QY 241 NASOHLTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQRLKLOTLMSSVDDSMETIYNN 300
DB 241 NASOHLTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQRLKLOTLMSSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYEFDIRVPPYVGRGPNVEAGCLNPHIV 360
DB 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYEFDIRVPPYVGRGPNVEAGCLNPHIV 360
QY 361 LNIDLAFTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDPSFLVERGKLL 420
DB 361 LNIDLAFTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDPSFLVERGKLL 420
QY 421 HKRDNDKVDQAQENFLPKYQVVKDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
DB 421 HKRDNDKVDQAQENFLPKYQVVKDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPBEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPBEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCYILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLREVGRHLLKKRPEEC 660
DB 601 HRCYILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLREVGRHLLKKRPEEC 660
QY 661 DCHKISYHTQHKGLKHRGSSLLHPRKGLQEKDKVLLREQKRLKLLKRLQNDTC 720
DB 661 DCHKISYHTQHKGLKHRGSSLLHPRKGLQEKDKVLLREQKRLKLLKRLQNDTC 720
QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCWRTINETHNPLFCFEFATGFL 780
DB 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCWRTINETHNPLFCFEFATGFL 780
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870
DB 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870
```

## RESULT 2

```
US-10-025-966A-6
; Sequence 6, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hemmerich
; TITLE OF INVENTION: Sulfatases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025, 966A
```

```
; CURRENT FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-025-966A-6
```

```
Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDQDVELGS 60
DB 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDQDVELGS 60
QY 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTKYVHNHNTYNNENCSFSSWA 120
DB 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTKYVHNHNTYNNENCSFSSWA 120
QY 121 QHESRTFAVINSTGYRTAFPGKYLNEYNGSVYPPGKKEWGLLKNSRFYNTLCRNGVK 180
DB 121 QHESRTFAVINSTGYRTAFPGKYLNEYNGSVYPPGKKEWGLLKNSRFYNTLCRNGVK 180
QY 181 EKHSYSDYLDLTLDLNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
DB 181 EKHSYSDYLDLTLDLNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
QY 241 NASOHLTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQRLKLOTLMSSVDDSMETIYNN 300
DB 241 NASOHLTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQRLKLOTLMSSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYEFDIRVPPYVGRGPNVEAGCLNPHIV 360
DB 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYEFDIRVPPYVGRGPNVEAGCLNPHIV 360
QY 361 LNIDLAFTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDPSFLVERGKLL 420
DB 361 LNIDLAFTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDPSFLVERGKLL 420
QY 421 HKRDNDKVDQAQENFLPKYQVVKDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
DB 421 HKRDNDKVDQAQENFLPKYQVVKDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPBEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPBEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCYILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLREVGRHLLKKRPEEC 660
DB 601 HRCYILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLREVGRHLLKKRPEEC 660
QY 661 DCHKISYHTQHKGLKHRGSSLLHPRKGLQEKDKVLLREQKRLKLLKRLQNDTC 720
DB 661 DCHKISYHTQHKGLKHRGSSLLHPRKGLQEKDKVLLREQKRLKLLKRLQNDTC 720
QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCWRTINETHNPLFCFEFATGFL 780
DB 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNNTYWCWRTINETHNPLFCFEFATGFL 780
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
```

QY 841 QYRQFRRKWPENKRPSSKSLGQLWEGWEG 870  
DB 841 QYRQFRRKWPENKRPSSKSLGQLWEGWEG 870  
RESULT 3  
US-10-025-966A-15  
; Sequence 15, Application US/10025966A  
; Publication No. US20030148920A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Steven D.  
; APPLICANT: Palmeri, Diana  
; APPLICANT: Stefan, Hemmerich  
; TITLE OF INVENTION: Sulfatases and methods of use  
; FILE REFERENCE: UCAL230  
; CURRENT APPLICATION NUMBER: US/10/025,966A  
; PRIOR FILING DATE: 2002-12-21  
; PRIOR APPLICATION NUMBER: 60/258,577  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 60/267,831  
; PRIOR FILING DATE: 2001-09-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-025-966A-15

Query Match 100.0%; Score 4750; DB 12; Length 870;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGPSLVLCLSATVFSLLGGSSAFSLSHRLKGRFDRDRNIRNIIILVLTDDQVELGS 60  
DB 1 MGPSLVLCLSATVFSLLGGSSAFSLSHRLKGRFDRDRNIRNIIILVLTDDQVELGS 60  
QY 61 MQVNNKTRIMEQGAHFINAFVTPMCCPSRSSILTKGVHNTYNNENCSPPSWQA 120  
DB 61 MQVNNKTRIMEQGAHFINAFVTPMCCPSRSSILTKGVHNTYNNENCSPPSWQA 120  
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRFFNYTLCRNGVK 180  
DB 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRFFNYTLCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGRFDRDRNIRNIIILVLTDDQVELGS 240  
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGRFDRDRNIRNIIILVLTDDQVELGS 240  
QY 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNN 300  
DB 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTIIVVTADHGHIHIGFGLVKSGMPEYEDIRVPFVYRGNVPEAGCLNPHIV 360  
DB 301 LVETGELDNTIIVVTADHGHIHIGFGLVKSGMPEYEDIRVPFVYRGNVPEAGCLNPHIV 360  
QY 361 LNIIDLAPTILDIAGLIDIPADMDGSIILKLDTERPVRNRFHLKGRFDRDRNIRNIIILVLTDDQVELGS 420  
DB 361 LNIIDLAPTILDIAGLIDIPADMDGSIILKLDTERPVRNRFHLKGRFDRDRNIRNIIILVLTDDQVELGS 420  
QY 421 HGRNDKVDAGEENFLPKYRQVDCQRAEYQTAACEQLGQKWCQVEDATCKLKLHCKKGP 480  
DB 421 HGRNDKVDAGEENFLPKYRQVDCQRAEYQTAACEQLGQKWCQVEDATCKLKLHCKKGP 480  
QY 481 MRLGSRALSNLVFKYCGGSEACTCGDYKLSLAGRRKGLFKKYKASVYRSIRSV 540  
DB 481 MRLGSRALSNLVFKYCGGSEACTCGDYKLSLAGRRKGLFKKYKASVYRSIRSV 540  
QY 541 AIEVDGRVYHVLGDAQAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAAPIKVT 600  
DB 541 AIEVDGRVYHVLGDAQAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAAPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKRPPSEC 660  
DB 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKRPPSEC 660  
QY 661 DCHKISYHTQHKGRGLKRGSSLHPPKGLQEKVKWLLRQKRRKKLRLQNNDTG 720  
DB 661 DCHKISYHTQHKGRGLKRGSSLHPPKGLQEKVKWLLRQKRRKKLRLQNNDTG 720  
QY 721 SMPGLTCFTHDNOHWQTAPFTWLTGPFCACTSANNNYWCMTINETHNFLFCEFAIGLE 780  
DB 721 SMPGLTCFTHDNOHWQTAPFTWLTGPFCACTSANNNYWCMTINETHNFLFCEFAIGLE 780  
QY 781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRNNDLGLKDGGSYE 840  
DB 781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRNNDLGLKDGGSYE 840  
QY 841 QYRQFRRKWPENKRPSSKSLGQLWEGWEG 870  
DB 841 QYRQFRRKWPENKRPSSKSLGQLWEGWEG 870

RESULT 4  
US-10-265-071-6  
; Sequence 6, Application US/10265071  
; Publication No. US20030147875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Steven D.  
; APPLICANT: Hemmerich, Stefan  
; APPLICANT: Tomita, Megumi  
; TITLE OF INVENTION: Sulfatases and methods of use thereof  
; FILE REFERENCE: UCAL-230CIP  
; CURRENT APPLICATION NUMBER: US/10/265,071  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,577  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 60/267,831  
; PRIOR FILING DATE: 2001-09-02  
; PRIOR APPLICATION NUMBER: 10/025,966  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-265-071-6

Query Match 100.0%; Score 4750; DB 12; Length 870;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGPSLVLCLSATVFSLLGGSSAFSLSHRLKGRFDRDRNIRNIIILVLTDDQVELGS 60  
DB 1 MGPSLVLCLSATVFSLLGGSSAFSLSHRLKGRFDRDRNIRNIIILVLTDDQVELGS 60  
QY 61 MQVNNKTRIMEQGAHFINAFVTPMCCPSRSSILTKGVHNTYNNENCSPPSWQA 120  
DB 61 MQVNNKTRIMEQGAHFINAFVTPMCCPSRSSILTKGVHNTYNNENCSPPSWQA 120  
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRFFNYTLCRNGVK 180  
DB 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRFFNYTLCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGRFDRDRNIRNIIILVLTDDQVELGS 240  
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGRFDRDRNIRNIIILVLTDDQVELGS 240  
QY 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNN 300  
DB 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTIIVVTADHGHIHIGFGLVKSGMPEYEDIRVPFVYRGNVPEAGCLNPHIV 360

Db 301 LVETGELDNITYIYVTTADHGVHIGQFGLVKGKSNPVEFDIRVPFYVRGPNVEAGCLNPHIV 360  
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSEFLVERGKLL 420  
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSEFLVERGKLL 420  
Qy 421 HKRDNDKVDQAQENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480  
Db 421 HKRDNDKVDQAQENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480  
Qy 481 MRLGGRSALSNLVPKYGGSEACTCDSDYKLSLAGREKLLFKKKYKASVYRSRSIRSV 540  
Db 481 MRLGGRSALSNLVPKYGGSEACTCDSDYKLSLAGREKLLFKKKYKASVYRSRSIRSV 540  
Qy 541 AIEVDGRVYHVLGDAQAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Db 541 AIEVDGRVYHVLGDAQAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Qy 601 HRCYIILENDTVQCDLILYKSLQAWKDKHLHIDHEITLQNKIKNREVRGHLKKRPEEC 660  
Db 601 HRCYIILENDTVQCDLILYKSLQAWKDKHLHIDHEITLQNKIKNREVRGHLKKRPEEC 660  
Qy 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKXKLLKRLKRLQNNDT 720  
Db 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKXKLLKRLKRLQNNDT 720  
Qy 721 SMPGLTCFTHDNOHWQTAFFWTGLGPFCACTSANNNTYWCRTINETHNLFCEPATGPLE 780  
Db 721 SMPGLTCFTHDNOHWQTAFFWTGLGPFCACTSANNNTYWCRTINETHNLFCEPATGPLE 780  
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840

RESULT 5

US-10-265-071-15  
; Sequence 15, Application US/10265071  
; Publication No. US20030147875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Steven D.  
; APPLICANT: Hammerich, Stefan  
; APPLICANT: Tomica, Megumi  
; TITLE OF INVENTION: Sulfatases and methods of use thereof  
; FILE REFERENCE: UCAL-230CIP  
; CURRENT APPLICATION NUMBER: US/10/265,071  
; PRIOR FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,577  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 60/267,831  
; PRIOR FILING DATE: 2001-09-02  
; PRIOR APPLICATION NUMBER: 10/025,966  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-265-071-15

Query Match 100.0%; Score 4750; DB 12; Length 870;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGPPSLVLCLSATVPSLGGSSAFSLHRLKGRFORRNIRPNIIILVLTDDQDVEIGS 60  
Db 1 MGPPSLVLCLSATVPSLGGSSAFSLHRLKGRFORRNIRPNIIILVLTDDQDVEIGS 60

Qy 61 MQVANKTRRIMEGGAHFINAFVTTMPCPSRSSILTKYVHNHNTYTNNECSSPSWQA 120  
Db 61 MQVANKTRRIMEGGAHFINAFVTTMPCPSRSSILTKYVHNHNTYTNNECSSPSWQA 120  
Qy 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWKWVGLLKNSRPFNYTLCRGVK 180  
Db 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWKWVGLLKNSRPFNYTLCRGVK 180  
Qy 181 EKSGSDYSKDYLTDLITNDSVSPFRSKMYPHRPVLWVISHAAPHGPDSDAPQYSLRFP 240  
Db 181 EKSGSDYSKDYLTDLITNDSVSPFRSKMYPHRPVLWVISHAAPHGPDSDAPQYSLRFP 240  
Qy 241 NASCHITPSTNYAPNDKHWIMEYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300  
Db 241 NASCHITPSTNYAPNDKHWIMEYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300  
Qy 301 LVETGELDNITYIYVTTADHGVHIGQFGLVKGKSNPVEFDIRVPFYVRGPNVEAGCLNPHIV 360  
Db 301 LVETGELDNITYIYVTTADHGVHIGQFGLVKGKSNPVEFDIRVPFYVRGPNVEAGCLNPHIV 360  
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSEFLVERGKLL 420  
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSEFLVERGKLL 420  
Qy 421 HKRDNDKVDQAQENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480  
Db 421 HKRDNDKVDQAQENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480  
Qy 481 MRLGGRSALSNLVPKYGGSEACTCDSDYKLSLAGREKLLFKKKYKASVYRSRSIRSV 540  
Db 481 MRLGGRSALSNLVPKYGGSEACTCDSDYKLSLAGREKLLFKKKYKASVYRSRSIRSV 540  
Qy 541 AIEVDGRVYHVLGDAQAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Db 541 AIEVDGRVYHVLGDAQAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Qy 601 HRCYIILENDTVQCDLILYKSLQAWKDKHLHIDHEITLQNKIKNREVRGHLKKRPEEC 660  
Db 601 HRCYIILENDTVQCDLILYKSLQAWKDKHLHIDHEITLQNKIKNREVRGHLKKRPEEC 660  
Qy 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKXKLLKRLKRLQNNDT 720  
Db 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKXKLLKRLKRLQNNDT 720  
Qy 721 SMPGLTCFTHDNOHWQTAFFWTGLGPFCACTSANNNTYWCRTINETHNLFCEPATGPLE 780  
Db 721 SMPGLTCFTHDNOHWQTAFFWTGLGPFCACTSANNNTYWCRTINETHNLFCEPATGPLE 780  
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Qy 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870  
Db 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870

RESULT 6

US-09-833-245-650  
; Sequence 650, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384

; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 650  
 ; LENGTH: 870  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-833-245-650

Query Match 100.0%; Score 4750; DB 12; Length 870;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFQDRRNIRNPIILVLTDDQDVELGS 60  
 DB 1 MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFQDRRNIRNPIILVLTDDQDVELGS 60  
 QY 61 MQVMKTRIMEOGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
 DB 61 MQVMKTRIMEOGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
 QY 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180  
 DB 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180  
 QY 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
 DB 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
 QY 241 NASQHITPSYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 DB 241 NASQHITPSYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 QY 301 LVEITGELDNITYIVYTADHGYHIGQFGLVKGSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
 DB 301 LVEITGELDNITYIVYTADHGYHIGQFGLVKGSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
 QY 361 LNIDLAPTILDIADIPADMDGKSLKLLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420  
 DB 361 LNIDLAPTILDIADIPADMDGKSLKLLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420  
 QY 421 HKRDNKVDQAQENFLPKYQVQKOLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480  
 DB 421 HKRDNKVDQAQENFLPKYQVQKOLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480  
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSIRSV 540  
 DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSIRSV 540  
 QY 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 DB 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 QY 601 HRCYILENDTVQCDDLDLYKSLQAWKDHKLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660  
 DB 601 HRCYILENDTVQCDDLDLYKSLQAWKDHKLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660  
 QY 661 DCHKISYHTQHKGRKLRGSSSLHPFRKGLQEKDKVLLRBEQFRKKLKLKRLQNNDTCTC 720  
 DB 661 DCHKISYHTQHKGRKLRGSSSLHPFRKGLQEKDKVLLRBEQFRKKLKLKRLQNNDTCTC 720  
 QY 721 SMPGLTCFTHDNQWQTAFTWTLGPPCACTSANNTYWCMTINETNINFLFCFATGFLF 780  
 DB 721 SMPGLTCFTHDNQWQTAFTWTLGPPCACTSANNTYWCMTINETNINFLFCFATGFLF 780  
 QY 781 YFDLNTDTPYQLMNAVNTLDRVLNQLHVLQMLRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 DB 781 YFDLNTDTPYQLMNAVNTLDRVLNQLHVLQMLRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 QY 841 QYRQFQRKWPMEKRPSSKSLGQLWEGWEG 870  
 DB 841 QYRQFQRKWPMEKRPSSKSLGQLWEGWEG 870

RESULT 7  
 US-09-970-287-2  
 ; Sequence 2, Application US/09970287  
 ; Publication No. US20030166283A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GLUCKSMANN, Maria A.  
 ; APPLICANT: MEYERS, Rachel  
 ; APPLICANT: KAPELLER-LIBERMANN, Rosana  
 ; APPLICANT: SILOS-SANTIAGO, Inmaculada  
 ; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR  
 ; FILE REFERENCE: 10147-61UI  
 ; CURRENT APPLICATION NUMBER: US/09/970,287  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: US 60/257,082  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 870  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-970-287-2

Query Match 99.9%; Score 4746; DB 12; Length 870;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFQDRRNIRNPIILVLTDDQDVELGS 60  
 DB 1 MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFQDRRNIRNPIILVLTDDQDVELGS 60  
 QY 61 MQVMKTRIMEOGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
 DB 61 MQVMKTRIMEOGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
 QY 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180  
 DB 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180  
 QY 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
 DB 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
 QY 241 NASQHITPSYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 DB 241 NASQHITPSYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 QY 301 LVEITGELDNITYIVYTADHGYHIGQFGLVKGSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
 DB 301 LVEITGELDNITYIVYTADHGYHIGQFGLVKGSMPEYFDIRVPFVYRGPVNEAGCLNPHIV 360  
 QY 361 LNIDLAPTILDIADIPADMDGKSLKLLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420  
 DB 361 LNIDLAPTILDIADIPADMDGKSLKLLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420  
 QY 421 HKRDNKVDQAQENFLPKYQVQKOLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480  
 DB 421 HKRDNKVDQAQENFLPKYQVQKOLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480  
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSIRSV 540  
 DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKCLFKKKYKASVYRSRSIRSV 540  
 QY 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 DB 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 QY 601 HRCYILENDTVQCDDLDLYKSLQAWKDHKLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660  
 DB 601 HRCYILENDTVQCDDLDLYKSLQAWKDHKLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660  
 QY 661 DCHKISYHTQHKGRKLRGSSSLHPFRKGLQEKDKVLLRBEQFRKKLKLKRLQNNDTCTC 720



Db 661 DCHKISYHTQHGKRLKRGSSLPFRKGLQEKDKVLLRQCKKKLRLQNNDTC 720  
QY 721 SMPGLTCTFTHDQWQTAPEFTLGPFCCTSANNTYKCMETINETHFLCFEATGFL 780  
Db 721 SMPGLTCTFTHDQWQTAPEFTLGPFCCTSANNTYKCMETINETHFLCFEATGFL 780  
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
QY 841 QYRQFORRWPEMKRPPSSKSLGQLWEGWEG 870  
Db 841 QYRQFORRWPEMKRPPSSKSLGQLWEGWEG 870

RESULT 8  
US-09-833-245-649  
; Sequence 649, Application US/09833245  
; Publication No. US200400101341  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 649  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (534)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-833-245-649

Query Match 99.9%; Score 4746; DB 12; Length 870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
Db 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
QY 61 MQVMNKTREIMEQGAHFINAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120  
Db 61 MQVMNKTREIMEQGAHFINAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120  
QY 121 QHESRTFAVLNLTGYRTAFFGKYLNEVNGSYVPPGKWEVGLLNKSRFYNTLCRNGVK 180  
Db 121 QHESRTFAVLNLTGYRTAFFGKYLNEVNGSYVPPGKWEVGLLNKSRFYNTLCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLTNDVSFFRTSKQYVPHRVLVISHAAPHGSDSAPQYSRLFP 240  
Db 181 EKHGSDYSKDYLTDLTNDVSFFRTSKQYVPHRVLVISHAAPHGSDSAPQYSRLFP 240  
QY 241 NASQHTPSYNTAPNDKHWIRYTGPMKPIHEWFTNMLQRKRLQTLMSVDDSMETIYNN 300  
Db 241 NASQHTPSYNTAPNDKHWIRYTGPMKPIHEWFTNMLQRKRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTIYVYADHGHYIQFGLVKGKMPYEFDIRVPFYRGPNVEAGCLNPHIV 360  
Db 301 LVETGELDNTIYVYADHGHYIQFGLVKGKMPYEFDIRVPFYRGPNVEAGCLNPHIV 360  
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSFLVERGKLL 420

Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRVWRDSFLVERGKLL 420  
QY 421 HKRDNDKVDQAQENFLPKYQVRKDI-CQRAEYQTAQCEQLGQKQCCVEDATGKLKHKCKGP 480  
Db 421 HKRDNDKVDQAQENFLPKYQVRKDI-CQRAEYQTAQCEQLGQKQCCVEDATGKLKHKCKGP 480  
QY 481 MRIGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540  
Db 481 MRIGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRSIRSV 540  
QY 541 ALEVGRVYHVGIGDAAPRNLTKEHWPGAPDQDDKGGDFSGTGGPLDYSAANPIKVT 600  
Db 541 ALEVGRVYHVGIGDAAPRNLTKEHWPGAPDQDDKGGDFSGTGGPLDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCCLDLVYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPPSEC 660  
Db 601 HRCYILENDTVQCCLDLVYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPPSEC 660  
QY 661 DCHKISYHTQHGKRLKRGSSLPFRKGLQEKDKVLLRQCKKKLRLQNNDTC 720  
Db 661 DCHKISYHTQHGKRLKRGSSLPFRKGLQEKDKVLLRQCKKKLRLQNNDTC 720  
QY 721 SMPGLTCTFTHDQWQTAPEFTLGPFCCTSANNTYKCMETINETHFLCFEATGFL 780  
Db 721 SMPGLTCTFTHDQWQTAPEFTLGPFCCTSANNTYKCMETINETHFLCFEATGFL 780  
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
QY 841 QYRQFORRWPEMKRPPSSKSLGQLWEGWEG 870  
Db 841 QYRQFORRWPEMKRPPSSKSLGQLWEGWEG 870

RESULT 9  
US-10-264-237-2754  
; Sequence 2754, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA131PI  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 2754  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (534)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-237-2754

Query Match 99.9%; Score 4746; DB 12; Length 870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
Db 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60  
QY 61 MQVMNKTREIMEQGAHFINAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120  
Db 61 MQVMNKTREIMEQGAHFINAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSQA 120

QY 121 QHESRTFAVILNSTGYRTAFKGYLNYNGSYVPPGKWEVGLLKNRFRFYNYTLCRNGVK 180  
DB 121 QHESRTFAVILNSTGYRTAFKGYLNYNGSYVPPGKWEVGLLKNRFRFYNYTLCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLITNDVSVPFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLFP 240  
DB 181 EKHGSDYSKDYLTDLITNDVSVPFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLFP 240  
QY 241 NASQHITPSYNAAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN 300  
DB 241 NASQHITPSYNAAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYFDIRVPPVYRGPVNEAGCLNPHIV 360  
DB 301 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYFDIRVPPVYRGPVNEAGCLNPHIV 360  
QY 361 LNIDLAPFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKMVRWDSFLVERGKLL 420  
DB 361 LNIDLAPFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKMVRWDSFLVERGKLL 420  
QY 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGKQWQCVEDATGKLLKHCKGCP 480  
DB 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGKQWQCVEDATGKLLKHCKGCP 480  
QY 481 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKLLFKKYYKASYVRSRSIRSV 540  
DB 481 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKLLFKKYYKASYVRSRSIRSV 540  
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
DB 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
QY 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLREVRGHLKKGPEEC 660  
DB 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLREVRGHLKKGPEEC 660  
QY 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNNTC 720  
DB 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNNTC 720  
QY 721 SMPGLTCFTHDNQHWQAPFTWTLGPPFCACTSANNNTYWCMTINETHNFLECFEATGFLE 780  
DB 721 SMPGLTCFTHDNQHWQAPFTWTLGPPFCACTSANNNTYWCMTINETHNFLECFEATGFLE 780  
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840  
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840  
QY 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870  
DB 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870

## RESULT 10

US-09-970-287-12

; Sequence 12, Application US/09970287

; Publication No. US2003016283A1

; GENERAL INFORMATION:

; APPLICANT: MEYERS, Rachel

; APPLICANT: KAPPELLER-LIBERMAN, Rosana

; APPLICANT: SILOS-SANTIAGO, Inmaculada

; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR

; FILE REFERENCE: 10147-61U1

; CURRENT APPLICATION NUMBER: US/09/970,287

; PRIOR FILING DATE: 2001-10-03

; PRIOR APPLICATION NUMBER: US 60/257,082

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12

; LENGTH: 885

TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-287-12

Query Match 99.9%; Score 4746; DB 12; Length 885;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPSLVLCILASATVFSLLGGSSAFLSHRLKGRFQDRRNRIRNIILVLTDDQDVELGS 60  
DB 16 MGPSLVLCILASATVFSLLGGSSAFLSHRLKGRFQDRRNRIRNIILVLTDDQDVELGS 75  
QY 61 MOVNKKTRIRIEQGGAFINAFVTPMCCPSRSILITGKYVHNENTYNNENCSPPWQA 120  
DB 76 MOVNKKTRIRIEQGGTHFINAFVTPMCCPSRSILITGKYVHNENTYNNENCSPPWQA 135  
QY 121 QHESRTFAVILNSTGYRTAFKGYLNYNGSYVPPGKWEVGLLKNRFRFYNYTLCRNGVK 180  
DB 136 QHESRTFAVILNSTGYRTAFKGYLNYNGSYVPPGKWEVGLLKNRFRFYNYTLCRNGVK 195  
QY 181 EKHGSDYSKDYLTDLITNDVSVPFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLFP 240  
DB 196 EKHGSDYSKDYLTDLITNDVSVPFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLFP 255  
QY 241 NASQHITPSYNAAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN 300  
DB 256 NASQHITPSYNAAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN 315  
QY 301 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYFDIRVPPVYRGPVNEAGCLNPHIV 360  
DB 316 LVETGELDNTYIVVTADHGYHIGQGLVKGKSMPEYFDIRVPPVYRGPVNEAGCLNPHIV 375  
QY 361 LNIDLAPFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKMVRWDSFLVERGKLL 420  
DB 376 LNIDLAPFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKMVRWDSFLVERGKLL 435  
QY 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGKQWQCVEDATGKLLKHCKGCP 480  
DB 436 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGKQWQCVEDATGKLLKHCKGCP 495  
QY 481 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKLLFKKYYKASYVRSRSIRSV 540  
DB 496 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKLLFKKYYKASYVRSRSIRSV 555  
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
DB 556 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 615  
QY 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLREVRGHLKKGPEEC 660  
DB 616 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLREVRGHLKKGPEEC 675  
QY 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNNTC 720  
DB 676 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNNTC 735  
QY 721 SMPGLTCFTHDNQHWQAPFTWTLGPPFCACTSANNNTYWCMTINETHNFLECFEATGFLE 780  
DB 736 SMPGLTCFTHDNQHWQAPFTWTLGPPFCACTSANNNTYWCMTINETHNFLECFEATGFLE 795  
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840  
DB 796 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 855  
QY 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870  
DB 856 QYRQFORRWKPEMKRPSKSLGQWEGWEG 885

## RESULT 11

US-09-989-722-84

; Sequence 84, Application US/09989722

; Patent No. US20020072067A1

GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tamas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C63  
CURRENT APPLICATION NUMBER: US/09/989,722  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.48; Score 4723.5; DB 9; Length 867;  
Best Local Similarity 99.78; Pred. No. 0;  
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRINRPNIILVLTDDQDVELGS 60  
DB 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRINRPNIILVLTDDQDVELGS 60  
QY 61 MOVNKKTRIMEOGGAHFINAFVTPMCCPSRSSILTKGVVHNHTYTNNECSPSWQA 120  
DB 61 MOVNKKTRIMEOGGAHFINAFVTPMCCPSRSSILTKGVVHNHTYTNNECSPSWQA 120  
QY 121 QHESRTFAVLNSTGYRTAFGKYLNEYNGSYVPPGKWKVGLLKNRFPNYTLCEGVK 180  
DB 121 QHESRTFAVLNSTGYRTAFGKYLNEYNGSYVPPGKWKVGLLKNRFPNYTLCEGVK 180  
QY 181 EKGSDYSKDYLTDLITNDSVSFFRSKQYPRHPVLWVISHAHPGSDSAPQYSRLFP 240  
DB 181 EKGSDYSKDYLTDLITNDSVSFFRSKQYPRHPVLWVISHAHPGSDSAPQYSRLFP 240

QY 241 NASOHITPSYNYAPNPDKHWIMRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNN 300  
DB 241 NASOHITPSYNYAPNPDKHWIMRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTIIVTADHGXYHIGQGLVKGKSMPEYEDIRVPFYVRGENVEAGCINPHIV 360  
DB 301 LVETGELDNTIIVTADHGXYHIGQGLVKGKSMPEYEDIRVPFYVRGENVEAGCINPHIV 360  
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVNRPHLKKQKRVWRDSFLVERGKLL 420  
DB 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVNRPHLKKQKRVWRDSFLVERGKLL 420  
QY 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYOTACEQLGQKQWQVEDATGKLLHKCKGP 480  
DB 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYOTACEQLGQKQWQVEDATGKLLHKCKGP 480  
QY 481 MELGGSRALSNLVPKYYGQGSSEACTCDSGDYKLSLAGRRKLLFKKCYKASYVRSRSIRSV 540  
DB 481 MELGGSRALSNLVPKYYGQGSSEACTCDSGDYKLSLAGRRKLLFKKCYKASYVRSRSIRSV 540  
QY 541 AIEVDGRVYHVLGDAAPRNLTGRHWPAGPEDDODDKGDFSGTGGGLPDYSAANPIKVT 600  
DB 541 AIEVDGRVYHVLGDAAPRNLTGRHWPAGPEDDODDKGDFSGTGGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHEIETLQNKIXNLREVGRHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHEIETLQNKIXNLREVGRHLKKRPEEC 660  
QY 661 DCHKISYHTQHKRLKHGSSLHPFRKGLQEKDKVLLRQKXKXKLLKRLQNNNDTC 720  
DB 661 DCHKISYHTQHKRLKHGSSLHPFRKGLQEKDKVLLRQKXKXKLLKRLQNNNDTC 720  
QY 721 SMPGLTCFTHDNQHWQAPFWTLGPRCACTSANNTYWCMTINETHNELECEBATGFLE 780  
DB 721 SMPGLTCFTHDNQHWQAPFWTLGPRCACTSANNTYWCMTINETHNELECEBATGFLE 780  
QY 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVQLMELFASCKYKQCNPRTNMDIQLKDGGSYE 840  
DB 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVQLMELFASCKYKQCNPRTNMDIQLKDGGSYE 840  
QY 841 QYRQFQRKWPMPKRPSSKSLGQLWEGWEG 870  
DB 838 QYRQFQRKWPMPKRPSSKSLGQLWEGWEG 867

RESULT 12  
US-09-989-723-84  
; Sequence 84, Application US/09989723  
; Patent No. US20020072092A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gertsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Tumanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C62  
CURRENT APPLICATION NUMBER: US/09/989,723  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25

Prior Application Number: 60/090694  
Prior Filing Date: 1998-06-25  
Prior Application Number: 60/090695  
Prior Filing Date: 1998-06-25  
Prior Application Number: 60/090696  
Prior Filing Date: 1998-06-25  
Prior Application Number: 60/090862  
Prior Filing Date: 1998-06-26  
Prior Application Number: 60/090863  
Prior Filing Date: 1998-06-26  
Prior Application Number: 60/091360  
Prior Filing Date: 1998-07-01  
Prior Application Number: 60/091478  
Prior Filing Date: 1998-07-02  
Prior Application Number: 60/091544  
Prior Filing Date: 1998-07-01  
Prior Application Number: 60/091519  
Prior Filing Date: 1998-07-02  
Prior Application Number: 60/091626  
Prior Filing Date: 1998-07-02  
Prior Application Number: 60/091633  
Prior Filing Date: 1998-07-02  
Prior Application Number: 60/091978  
Prior Filing Date: 1998-07-07  
Prior Application Number: 60/091982  
Prior Filing Date: 1998-07-07  
Prior Application Number: 60/092182  
Prior Filing Date: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 9; Length 867;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

2y 1 MGPPSLVLCSSATVPSLIGGSAFSLHRLKGRFQDRDRNIRPNILVLTDDQVELGS 60  
2b 1 MGPPSLVLCSSATVPSLIGGSAFSLHRLKGRFQDRDRNIRPNILVLTDDQVELGS 60  
2y 61 MQVWNTKTRIMEOGGAFFNAFVTTDMCCPSRSSILTGKVVHNHTYTNENCSFSPQA 120  
2b 61 MQVWNTKTRIMEOGGAFFNAFVTTDMCCPSRSSILTGKVVHNHTYTNENCSFSPQA 120  
2y 121 QHSRTFAVNLSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLNKSNRYNTTLCRNGVK 180  
2b 121 QHSRTFAVNLSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLNKSNRYNTTLCRNGVK 180  
2y 181 EKHGSDYSKDYLTDLTNDVSFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSRLFP 240  
2b 181 EKHGSDYSKDYLTDLTNDVSFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSRLFP 240  
2y 241 NASCHITPSYNAFNPDKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
2b 241 NASCHITPSYNAFNPDKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
2y 301 LVETGELDNTYIYVTADHGHVIGQGLVKGKSNPYEFDIRVPFYVGRGNVAGCLNPHIV 360  
2b 301 LVETGELDNTYIYVTADHGHVIGQGLVKGKSNPYEFDIRVPFYVGRGNVAGCLNPHIV 360  
2y 361 LNIDLAPTILDIADLIPADMDGKSLIKLIDTFRPVNRPHLKKQKVRVDRSFLVBERGKLL 420  
2b 361 LNIDLAPTILDIADLIPADMDGKSLIKLIDTFRPVNRPHLKKQKVRVDRSFLVBERGKLL 420  
2y 421 HKRDNKVDQAEENFLPKYQVRVXDLQRAEYQACBOLGQKQVCVEDATGKLLKHKCKGP 480  
2b 421 HKRDNKVDQAEENFLPKYQVRVXDLQRAEYQACBOLGQKQVCVEDATGKLLKHKCKGP 480  
2y 481 MRLGGSRALSNLVPKYQGQSEACTCDSDGYKLSLAGRKKLPKKYKASYVRSRSIRSV 540  
2b 481 MRLGGSRALSNLVPKYQGQSEACTCDSDGYKLSLAGRKKLPKKYKASYVRSRSIRSV 540  
2y 541 AIEVDGRVHVGLGDAAPQPNLTKRHWPAGPEQDDKQGGDFSGTGGLPDYSAANPIKVT 600  
2b 541 AIEVDGRVHVGLGDAAPQPNLTKRHWPAGPEQDDKQGGDFSGTGGLPDYSAANPIKVT 600

Qy 601 HRCVILENTVQCDLDBLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKKQPEEC 660  
Db 601 HRCVILENTVQCDLDBLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKKQPEEC 660  
Qy 661 DCHKISVHTQKHGRLKHRRGSSLHPFRKGLQEKDKVLLREQKKKLLKLLKELQNDTC 720  
Db 661 DCHKISVHTQKHGRLKHRRGSSLHPFRKGLQEKDKVLLREQKKKLLKLLKELQNDTC 720  
Qy 721 SMPGLTCFTHDNQHWOTAPFWTLG2PFCACTSANNNTYWCMTINETHNLFCEFBATGFE 780  
Db 721 SMPGLTCFTHDNQHWOTAPFWTLG2PFCACTSANNNTYWCMTINETHNLFCEFBATGFE 780  
Qy 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVQLWEURSCGYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVQLWEURSCGYKQCNPRTRNMDLGLKDGGSYE 840  
Qy 841 QYRQFQRKWPENKMPSSKSLGQWEGWEG 870  
Db 838 QYRQFQRKWPENKMPSSKSLGQWEGWEG 867

RESULT 13

US-09-989-279-84

; Sequence 84, Application US/09989279

; Patent No. US20020072496A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kijavini, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2730P1C56

; CURRENT APPLICATION NUMBER: US/09/989,279

; CURRENT FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07



PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089801
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090245
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090252
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090254
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090343
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090355
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090429
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090431
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090435
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090444
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090445
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090542
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090557
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090676
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090678
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090690
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090694
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090695
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090696
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090862
PRIOR FILING DATE:	1998-06-26
PRIOR APPLICATION NUMBER:	60/090863
PRIOR FILING DATE:	1998-06-26
PRIOR APPLICATION NUMBER:	60/091360
PRIOR FILING DATE:	1998-07-01
PRIOR APPLICATION NUMBER:	60/091478
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091544
PRIOR FILING DATE:	1998-07-01
PRIOR APPLICATION NUMBER:	60/091519
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091626
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091633
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091978
PRIOR FILING DATE:	1998-07-01

PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match	99.4%	Score 4723.5	DB 9	Length 867
Best Local Similarity	99.7%	Pred. No. 0	Mismatches 0	Indels 3
Matches 867	Conservative 0			Gaps 1
1	MGPPSLVLCCLLSATVFLSGLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS	60		
1	MGPPSLVLCCLLSATVFLSGLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS	60		
61	MOVNKRTRIMEGGGAHFNATVTPMCCPSRSSILTCKYVHNHTYNNENCSPSQA	120		
61	MOVNKRTRIMEGGGAHFNATVTPMCCPSRSSILTCKYVHNHTYNNENCSPSQA	120		
121	QHSRTFAVYLNSTGYRTAFFGKYLNEYNGSVPPGKWEVGLLKNRSFYNYTLCRNGVK	180		
121	QHSRTFAVYLNSTGYRTAFFGKYLNEYNGSVPPGKWEVGLLKNRSFYNYTLCRNGVK	180		
181	EXHGSYSDYITDILTNDSVFFRTSKOMYHPRLVLMVISHAAPHGPEDSAPQVSRLLFP	240		
181	EXHGSYSDYITDILTNDSVFFRTSKOMYHPRLVLMVISHAAPHGPEDSAPQVSRLLFP	240		
241	NASOHITPSYNYAPNPKDWIRYTGPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN	300		
241	NASOHITPSYNYAPNPKDWIRYTGPKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN	300		
301	LVETGELDNITYVYTADHGYYH:GQFGLVKGKSMPEYDIPVFPVYVRGPNVAGCLNPHIV	360		
301	LVETGELDNITYVYTADHGYYH:GQFGLVKGKSMPEYDIPVFPVYVRGPNVAGCLNPHIV	360		
361	LNIDLAPTITLDIAGLDIPADMDGKSLTKLLDTERPVNRPFLKXKRWVRDSDLVERGKLL	420		
361	LNIDLAPTITLDIAGLDIPADMDGKSLTKLLDTERPVNRPFLKXKRWVRDSDLVERGKLL	420		
421	HKRDNKDVAQENFLPKYQRVKDLQRAEYQACEQLGQKQWQCVEDATGKLLHKCKGP	480		
421	HKRDNKDVAQENFLPKYQRVKDLQRAEYQACEQLGQKQWQCVEDATGKLLHKCKGP	480		
481	MRLGGSRALSNLVPKYGGGSACTCDSDGYKLSLAGRRKKLFKKKYKASVYRSRSIRSV	540		
481	MRLGGSRALSNLVPKYGGGSACTCDSDGYKLSLAGRRKKLFKKKYKASVYRSRSIRSV	540		
541	ALIEVDGRVHVGLGDAAPRNLTCKHWPAPEDQDDKGGDFSGTGGLPDYSAANPIKVT	600		
541	ALIEVDGRVHVGLGDAAPRNLTCKHWPAPEDQDDKGGDFSGTGGLPDYSAANPIKVT	600		
601	HRCYILENDTVQCDLDLYKSLQAWKDLHLIDHIEITLQNKIKNLREVGRHLKKGKPEEC	660		
601	HRCYILENDTVQCDLDLYKSLQAWKDLHLIDHIEITLQNKIKNLREVGRHLKKGKPEEC	660		
661	DCHKISYHTQHKRLKRGSSLLHPRKGLQEKDKWLLREOKRKKLKLKRLQNNDTIC	720		
661	DCHKISYHTQHKRLKRGSSLLHPRKGLQEKDKWLLREOKRKKLKLKRLQNNDTIC	720		
721	SMPLGTCFTHDNQHWQAPFWTLGPFCACTSANNNTYWCWRTINETHNLFCEFAFGLE	780		
721	SMPLGTCFTHDNQHWQAPFWTLGPFCACTSANNNTYWCWRTINETHNLFCEFAFGLE	780		
781	YFDLNTDPTQLMNAVNTLDRDVLNQLHVQMLRSCGKYQCNQPNTRMDLGLKDGGSYE	840		
781	YFDLNTDPTQLMNAVNTLDRDVLNQLHVQMLRSCGKYQCNQPNTRMDLGLKDGGSYE	840		
841	QYRQFORRWKPMKRPSSKSLGLQWEGWG 870			
838	QYRQFORRWKPMKRPSSKSLGLQWEGWG 867			

RESULT 14  
US-09-989-727-84  
; Sequence 84, Application US/09989727

Patent No. US20020072497A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoli, Nicholas P.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC65  
CURRENT APPLICATION NUMBER: US/09/989,727  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 9; Length 867;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLLKGRFQDRRNIRNIIILVLTDDDDVELGS 60  
Db 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLLKGRFQDRRNIRNIIILVLTDDDDVELGS 60  
QY 61 MQVMNKTTRIMEQGAHFNAFVTPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
Db 61 MQVMNKTTRIMEQGAHFNAFVTPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120  
QY 121 QHESRTFAVLNLTGRTAFKGYLNEVNGSVPPGKWKVLLKNSRFYNTILCRNGVK 180  
Db 121 QHESRTFAVLNLTGRTAFKGYLNEVNGSVPPGKWKVLLKNSRFYNTILCRNGVK 180  
QY 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPYSRLFP 240  
Db 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPYSRLFP 240

241 NASQHTPSYNTAPNPKWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
241 NASQHTPSYNTAPNPKWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
301 LVETGELDNITYIVYTADHGVHIGQFGLVKGKSMPEYFDIRVFPYVRGNVAGCLNPHIV 360  
301 LVETGELDNITYIVYTADHGVHIGQFGLVKGKSMPEYFDIRVFPYVRGNVAGCLNPHIV 360  
361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKKRVWRDLSFLVERGKLL 420  
361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKKRVWRDLSFLVERGKLL 420  
421 HKENDNDKVAQENFUPKYQVVKDLCORAEYQACEQLGQKQWOCVEDATGKLLHKCKGP 480  
421 HKENDNDKVAQENFUPKYQVVKDLCORAEYQACEQLGQKQWOCVEDATGKLLHKCKGP 480  
481 MRLGGSRALSNLVPKYGGSEACTCSDGYKLSLAGREKKLFKKYKASVYRSRSISV 540  
481 MRLGGSRALSNLVPKYGGSEACTCSDGYKLSLAGREKKLFKKYKASVYRSRSISV 540  
541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDODDKGGDFSGTGLPDYSAANPIKVT 600  
541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDODDKGGDFSGTGLPDYSAANPIKVT 600  
601 HRCYILENDTVQCDLKYSLQAKDKHLHIDHIEITLQNKILNREVGRHLKKRPREC 660  
601 HRCYILENDTVQCDLKYSLQAKDKHLHIDHIEITLQNKILNREVGRHLKKRPREC 660  
661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVWLLREQKRRKKLRLKRLQNNDT 720  
661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVWLLREQKRRKKLRLKRLQNNDT 720  
721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCWRTINETHNPLFCBPATGFILE 780  
721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCWRTINETHNPLFCBPATGFILE 780  
781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVLQMLRSCGYKQCNPRTRNMDLKGDDGSYE 840  
781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVLQMLRSCGYKQCNPRTRNMDLKGDDGSYE 840  
841 OYRQFORRWKPEKRPSSKSLQOLWEGWEG 870  
838 OYRQFORRWKPEKRPSSKSLQOLWEGWEG 867

RESULT 15

JS-09-989-731-84

; Sequence 84, Application US/09989731

; Patent No. US20020103125A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C70  
; CURRENT APPLICATION NUMBER: US/09/989,731  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10

1	PRIOR APPLICATION NUMBER: 60/088855
2	PRIOR FILING DATE: 1998-06-11
3	PRIOR APPLICATION NUMBER: 60/088861
4	PRIOR FILING DATE: 1998-06-11
5	PRIOR APPLICATION NUMBER: 60/088876
6	PRIOR FILING DATE: 1998-06-11
7	PRIOR APPLICATION NUMBER: 60/089105
8	PRIOR FILING DATE: 1998-06-12
9	PRIOR APPLICATION NUMBER: 60/089440
10	PRIOR FILING DATE: 1998-06-16
11	PRIOR APPLICATION NUMBER: 60/089512
12	PRIOR FILING DATE: 1998-06-16
13	PRIOR APPLICATION NUMBER: 60/089514
14	PRIOR FILING DATE: 1998-06-16
15	PRIOR APPLICATION NUMBER: 60/089532
16	PRIOR FILING DATE: 1998-06-17
17	PRIOR APPLICATION NUMBER: 60/089538
18	PRIOR FILING DATE: 1998-06-17
19	PRIOR APPLICATION NUMBER: 60/089598
20	PRIOR FILING DATE: 1998-06-17
21	PRIOR APPLICATION NUMBER: 60/089599
22	PRIOR FILING DATE: 1998-06-17
23	PRIOR APPLICATION NUMBER: 60/089600
24	PRIOR FILING DATE: 1998-06-17
25	PRIOR APPLICATION NUMBER: 60/089653
26	PRIOR FILING DATE: 1998-06-17
27	PRIOR APPLICATION NUMBER: 60/089801
28	PRIOR FILING DATE: 1998-06-18
29	PRIOR APPLICATION NUMBER: 60/089907
30	PRIOR FILING DATE: 1998-06-18
31	PRIOR APPLICATION NUMBER: 60/089908
32	PRIOR FILING DATE: 1998-06-18
33	PRIOR APPLICATION NUMBER: 60/089947
34	PRIOR FILING DATE: 1998-06-19
35	PRIOR APPLICATION NUMBER: 60/089948
36	PRIOR FILING DATE: 1998-06-19
37	PRIOR APPLICATION NUMBER: 60/089952
38	PRIOR FILING DATE: 1998-06-19
39	PRIOR APPLICATION NUMBER: 60/090246
40	PRIOR FILING DATE: 1998-06-22
41	PRIOR APPLICATION NUMBER: 60/090252
42	PRIOR FILING DATE: 1998-06-22
43	PRIOR APPLICATION NUMBER: 60/090254
44	PRIOR FILING DATE: 1998-06-22
45	PRIOR APPLICATION NUMBER: 60/090349
46	PRIOR FILING DATE: 1998-06-23
47	PRIOR APPLICATION NUMBER: 60/090355
48	PRIOR FILING DATE: 1998-06-23
49	PRIOR APPLICATION NUMBER: 60/090429
50	PRIOR FILING DATE: 1998-06-24
51	PRIOR APPLICATION NUMBER: 60/090431
52	PRIOR FILING DATE: 1998-06-24
53	PRIOR APPLICATION NUMBER: 60/090435
54	PRIOR FILING DATE: 1998-06-24
55	PRIOR APPLICATION NUMBER: 60/090557
56	PRIOR FILING DATE: 1998-06-24
57	PRIOR APPLICATION NUMBER: 60/090535
58	PRIOR FILING DATE: 1998-06-24
59	PRIOR APPLICATION NUMBER: 60/090445
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090472
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090557
64	PRIOR FILING DATE: 1998-06-24
65	PRIOR APPLICATION NUMBER: 60/090676
66	PRIOR FILING DATE: 1998-06-25
67	PRIOR APPLICATION NUMBER: 60/090678
68	PRIOR FILING DATE: 1998-06-25
69	PRIOR APPLICATION NUMBER: 60/090690
70	PRIOR FILING DATE: 1998-06-25

[illegible]

2Y 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 661 DCHKISYHTQHKGRHLKRGSSSLHPPFRKGLQEKDKVWLLREQRKKKRLKRLQNNDT 720  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 661 DCHKISYHTQHKGRHLKRGSSSLHPPFRKGLQEKDKVWLLREQRKKKRLKRLQNNDT 720  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 721 SMPGLTCFTHDNOHWOTAPFTWLGPFCACTSANNTYWCMTINETHNLFCEFATGFE 780  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 721 SMPGLTCFTHDNOHWOTAPFTWLGPFCACTSANNTYWCMTINETHNLFCEFATGFE 780  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 781 YFDLNTDPYQLMNAVNTLDRDVLNQHLVQLMELRSCGYKQCNPRTRNDLGLKDGGSYE 840  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 781 YFDLNTDPYQLMNAVNTLDRDVLNQHLVQLMELRSCGYKQCNPRTRNDLGLKDGGSYE 840  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 841 QYRQFORRKWPEMKRPSSKSLGQWEGWEG 870  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2Y 838 QYRQFORRKWPEMKRPSSKSLGQWEGWEG 867  
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: February 14, 2004, 23:07:49  
Job time : 83 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

Run on: February 14, 2004, 19:42:15 ; Search time 39 seconds  
(without alignments)  
1049.057 Million cell updates/sec

Title: US-10-025-966A-6  
Effect score: 4750  
Sequence: 1 MGPPSLVLCLSATVFSLLG.....PEMKRPSKSLQWEGWEG 870

Coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	858.5	18.1	552	1	GL6S HUMAN	P15586 homo sapien
2	856	18.0	559	1	GL6S_CAPHI	P50426 capra hircu
3	452.5	9.5	649	1	ARS_VOLCA	Q10723 volvox cart
4	405	8.5	647	1	ARS_CHIRE	P14217 chlamydomon
5	302.5	6.4	497	1	YIDJ_ECOLI	P31447 escherichia
6	250	5.5	512	1	BETC_RHIME	O69787 rhizobium m
7	251	5.3	560	1	YDSE_ECOLI	P77318 escherichia
8	238.5	5.0	583	1	STS_HUMAN	P08842 homo sapien
9	238	5.0	567	1	ARS_STRPU	P50473 strongyloce
10	234.5	4.9	550	1	IDS_HUMAN	P22304 homo sapien
11	232	4.9	551	1	ARS_HEMPU	P14000 hemientrot
12	226	4.8	563	1	IDS_MOUSE	Q08890 mus musculu
13	218.5	4.6	522	1	GAGS_HUMAN	P34059 homo sapien
14	210.5	4.4	506	1	ARS_MOUSE	P50428 mus musculu
15	210.5	4.4	533	1	ARS_HUMAN	P15848 homo sapien
16	210.5	4.4	577	1	STS_RAT	P15589 rattus norv
17	203	4.3	535	1	ARS_FELCA	P33727 felis silve
18	201.5	4.2	464	1	ARS_KLEAE	P20713 klebsiella
19	200	4.2	593	1	ARSD_HUMAN	P51689 homo sapien
20	198.5	4.2	591	1	ARSF_HUMAN	P54793 homo sapien
21	194.5	4.1	507	1	ARSA_HUMAN	P15289 homo sapien
22	189.5	4.0	624	1	STS_MOUSE	P50427 mus musculu
23	183.5	3.9	589	1	ARSE_HUMAN	P51690 homo sapien
24	176.5	3.7	502	1	SPHM_HUMAN	P51688 homo sapien
25	176	3.7	551	1	ASLA_ECOLI	P25549 escherichia
26	170	3.6	473	1	ARS_MOUSE	P50430 rattus norv
27	155.5	3.3	535	1	ARS_PSEAE	P51691 pseudomonas
28	138.5	2.9	510	1	GPNI_LISIN	Q92812 listeria in
29	137.5	2.9	510	1	GPNI_LISMO	Q8Y414 listeria mo
30	135.5	2.9	510	1	GPNI_LISMO	Q8Y414 listeria mo
31	134	2.8	533	1	GPNI_ANTSP	Q05464 antithamio
32	128	2.7	285	1	ARSB_MOUSE	Q8Y412 anabaena sp
33	127	2.7	532	1	GPNI_SYNY3	P74507 synecocyst

#### RESULT 1

GL6S_HUMAN	2.6	724	1	AT11_VACCV	P24759 vaccinia vi
ID	124	510	1	GPML_BACHD	Q9k716 bacillus ba
AC	124	512	1	GPML_OCEIH	P59174 oceanobacil
DT	123	2339	1	RPCI_PLAFA	P27625 plasmodium
DT	122	510	1	GPML_CLOAB	Q97153 clostridium
38	121	534	1	GPML_PORPU	P51379 porphyra pu
39	121	534	1	GPML_PORPU	Q13206 homo sapien
40	118.5	875	1	DD10_HUMAN	Q9x519 bacillus st
41	118	510	1	GPML_BACST	P54496 bacillus su
42	118	510	1	YQGS_BACSU	Q8XK22 clostridium
43	117	512	1	GPML_CLOPE	P55155 caenorhabdi
44	112	1616	1	VITI_CABEL	P19214 plasmodium
45	111.5	2.3	1435	EBAL_PLAFC	

#### ALIGNMENTS

GL6S_HUMAN	STANDARD;	PRT;	552 AA.
AC	P15586;		
DT	01-APR-1990 (Rel. 14, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	N-acetylglucosamine-6-sulfatase precursor (EC 3.1.6.14) (G6S)		
DE	(Glucosamine-6-sulfatase).		
GN	GNS.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID:9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-Endothelial cells;		
RX	MEDLINE-9309807; PubMed:1463457;		
RA	Robertson D.A., Freeman C., Morris C.P., Hopwood J.J.;		
RT	"A cDNA clone for human glucosamine-6-sulfatase reveals differences between arylsulphatases and non-arylsulphatases.";		
RL	Biochem. J. 288:539-544 (1992).		
RP	SEQUENCE OF 178-552 FROM N.A., AND PARTIAL SEQUENCE.		
RX	MEDLINE-89061714; PubMed:3196333;		
RA	Robertson D.A., Freeman C., Nelson P.V., Morris C.P., Hopwood J.J.;		
RT	"Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase.";		
RL	Biochem. Biophys. Res. Commun. 157:218-224 (1988).		
CC	- - CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate group of the N-acetyl-D-glucosamine 6-sulfate units of heparan sulfate and keratan sulfate.		
CC	- - SUBCELLULAR LOCATION: Lysosomal.		
CC	- - DISEASE: DEFECTS IN GNS RESULT IN THE ACCUMULATION OF PARTIALLY DEGRADED HEPARAN SULFATE IN LYSOSOMES CAUSING ORGANELLE, CELL AND TISSUE DISTORTION, ULTIMATELY LEADING TO THE LYSOSOMAL STORAGE DISORDER, MUCOPOLYSACCHARIDOSIS TYPE IIID (MPS-IIID) (ALSO KNOWN AS SANFILIPPO D SYNDROME).		
CC	- - SIMILARITY: BELONGS TO THE SULFATASE FAMILY.		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	EMBL; Z12173; CAA78164.1; --		
DR	PIR; S27164; KKHUGU.		
DR	Genew; HGNC:4422; GNS.		
DR	MIM; 252940; --		
DR	GO; GO:0008449; F:N-acetylglucosamine-6-sulfatase activity; TAS.		
DR	GO; GO:0006027; P:glycosaminoglycan catabolism; TAS.		
DR	InterPro; IPR000917; Sulfatase.		
DR	Fram; PF00884; Sulfatase; 1.		



T CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).  
T CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).  
Q SEQUENCE 559 AA; 62711 MW; 786CDDC48334A458 CRC64;  
Query Match 18.0%; Score 856; DB 1; Length 559;  
Best Local Similarity 27.5%; Pred. NO. 2.18-54;  
Matches 231; Conservative 77; Mismatches 177; Indels 354; Gaps 20;  
Y 3 PPSVLCLLSATVFLSLLGSSAFSLSHRLKGRFORDRNRIRPNILVLTDQDQVELGSMQ 62  
b 29 PPLPLLL-----LGGCLG-----VSGAAGSR---RPNVVLVLAADQDEVLGMT 72  
Y 63 VNKTTRIMEQGGAHFNATVTPMCCPSRSSILTGKIVHNHNTYTN--NENCSSPSQWA 120  
b 73 PLKTKALIGEMGMTFSSAYVPSALCCPSRASILTGKYPHNHVNNTLEGCSSKSWQK 132  
Y 121 QHESRTRAVLNS--TGRTAFPGKLYNNG-----SYVPPGKWEVGLLKNRFRNYT 173  
b 133 IQEPNTFFAILRSMCGYQTFAGKLYNBYGAPDAGGLGHVPLGHSYVALEKSKYNYT 192  
Y 174 LCRNGVKRKGSDYSKDYLTDLITNDSVP--FRTSKMYPHRPVLMVISHAAPHGPEDS 231  
b 193 LSLNGKARKHGENYSDVLTDLVLANVSLDFLYKNSN-----PFFMMISTPAPHSPTA 247  
Y 232 APOYSRLPNASQHTPSVYNAPNDKHWIMRYT--GPMKPIHMBFTNMQEKLQTLMSV 290  
b 248 APOYQNAFONFAPRNKNFN--IHGTNKHILRQAKTPMNSIOFLNAPFRWQTLTSLV 306  
Y 291 DDSMETIYNMLVETGELDNTVYVYADHYHIGQGLVKGKSMPEYEDIRVFFVVRGPNV 350  
b 307 DDLVEKLKRLFEFNGELNNTVYFYSNGYHTGQPSLPIDKRLQYEFEDIKVPVLVRGPGI 366  
Y 351 EAGLNPVHVLNIDAPLIDTILAGLDI--PADMDGSKILKILDTVPYRFLKXKRVMR 409  
b 367 KPNQTSKMLVANIDLPITLIDTAGYGLNKTQMDGMSFLPIL---RGASNL-----TWR 416  
Y 410 DFLYVERGKLLHKRNDKVDQAENFLPKYQKVDLCQABYQYQACEQLGQKQWQVEDAT 469  
b 417 SDVLVE-----RNV----- 422  
Y 470 GKLLHKCKGPMRLGGSRALSNLPKYVQGSSEACTDSGDYKLSLAGRRKKLPKKYKA 529  
b 423 -----YQEG----- 427  
Y 530 SYVRSRSIRSALEVDGRVYHVLGDAAQPNLTKRHWPGAPEDQDDGDFSGTGGLP 589  
b 428 -----RNV----- 431  
Y 590 DYSAANPIKVRHCYILENTVQCDDLYKSLQAWKHLHIDHETLQNKIKNLREVR 649  
b 432 ----- 431  
Y 650 GHLKKRPECDCXKTSYTHQKGLKXGSSLHPRKGLQEKDKVLLRQKXKXKLK 709  
b 432 -----DPTCPSL----- 439  
Y 710 LAKRLQNDTCSMPGLT--CFTHDNQHWOTAPFWTLPGFCACTSANNNTYCWMTINETHN 768  
b 440 -----PGVSCF-----PDCVCEDAYNTYACVRTMSELWN 470  
Y 769 FLFCEP--ATGFELEPDINTDYPQLMAVNTLDRDVLNQLHVLMLSCGKYKOCNPR 825  
b 471 LOYCEPDQDEVEFVYVNTADPHQLNNAKSIDPELLGQNYRLMLQSCSGFTCTRTPR 529

RESULT 3  
ABS\_VOLCA STANDARD; PRT; 649 AA.  
ID AC Q10723;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).

OS Volvox carteri.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Volvocaceae; Volvox.  
OX NCBI\_TaxID=3067;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN=f. Nagariensis / HK10;  
RX MEDLINE=94222055; PubMed=8168504;  
RA Hallmann A., Sumper M.;  
RT "An inducible arylsulfatase of Volvox carteri with properties  
RT suitable for a reporter-gene system. Purification, characterization  
RT and molecular cloning";  
RL Eur. J. Biochem. 221:143-150(1994).  
RN [2]  
RP SEQUENCE OF 64-76, AND POST-TRANSLATIONAL MODIFICATION OF CYG-72.  
RX MEDLINE=96283826; PubMed=8681943;  
RA Selmer T., Hallmann A., Schmidt B., Sumper M., von Figura K.;  
RT "The evolutionary conservation of a novel protein modification, the  
RT conversion of cysteine to serinesemialdehyde in arylsulfatase from  
RT Volvox carteri";  
RL Eur. J. Biochem. 238:341-345(1996).  
CC -I- FUNCTION: Is commonly produced by soil microorganisms and plays an  
CC important role in the mineralization of sulfates.  
CC -I- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.  
CC -I- ENZYME REGULATION: INHIBITED BY NA(3)BO(3) AND KCN. NO INHIBITION  
CC BY SODIUM DODECYL SULFATE, EVEN AT HIGH CONCENTRATION.  
CC -I- SUBCELLULAR LOCATION: Periplasmic.  
CC -I- INDUCTION: By sulfur deprivation.  
CC -I- MISCELLANEOUS: The enzyme is thermostable, exhibiting a  
CC temperature optimum at 60 degrees Celsius. Its optimal pH is 8.0.  
CC -I- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC ENBL; X77214; CAA54426.1; -;  
DR PIR; S43229; S43229.  
DR InterPro; IPR000917; Sulfatase.  
DR Pfam; PF00884; Sulfatase; 1.  
DR PROSITE; PS00523; SULFATASE\_1; 1.  
DR PROSITE; PS00149; SULFATASE\_2; 1.  
KW Hydrolase; Signal; Glycoprotein; Periplasmic.  
FT SIGNAL 1 22  
FT CHAIN 23 649 ARYLSULFATASE.  
FT MOD RES 72 72 2-AMINO-3-OXOPROPIONIC ACID.  
FT DOMAIN 571 582 POLY-ALA.  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 649 AA; 72287 MW; 0C23EFD77C43F7B9 CRC64;

Query Match 9.5%; Score 452.5; DB 1; Length 649;  
Best Local Similarity 27.4%; Pred. NO. 4.6e-25;  
Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;  
QY 6 LVLCLLSATVFLSLLGSSAFSLSHRLKGRFORDRNRIRPNILVLTDQDQVELGSM--QV 63  
DB 7 VALCLLG---FAALTAAA-----HQ-----RPNFVVFITDDQDQGINSTHPY 47  
QY 64 MNKTRIMEQGGAHFNATVTPMCCPSRSSILTGKIVHNHNTYTNENCSSP-----SW 118  
DB 48 QPKLHEHRIYFGLHKLKNYFVTTPVCCPSRLNWRGQFSN-----TNFTDVLGPHGYSKW 103  
QY 119 QAQHSRTF--AVYLNSTGVRTAFPGKYLNEY---NGSYVPPGKWEVGLLKNRFR--YNYT 173

```
Db 104 KSLGIDKSYLWNLQNGYNTYVVGKFLVDYSVSNQYQNPAGWTDIDALVTPYTDYNNP 163
2y 174 -LCRNGVKEKGGSDYKDYLTDLITNDSVFPRTSKMYPHRPVLMWISHAAPHGPDSA 232
Db 164 GFSRNGATP---NIYPGFYSTVDIAKVAQAKTA--VAAGKFFYAQISPIAPH-----212
2y 233 PQYSRLFPNASQHITPSYNYAPNP-DKHWIMRYTGP-----KPIHMEF 275
Db 213 -TSTQIVFDVANATKFFYPPIAPRHWELFSDATLPGTSHKNLYEADVSDKAWIRA 271
2y 276 TNNLQ-----RKRLQTLMSVDDSMETINMLVETGELDNITYVYTDHGHVHGG 324
Db 272 LPLAQNNRTYLEVYRLRLSLASDELIDRVVATLQEGVLDNTYLYISADNGYHVGT 331
2y 325 FGLVKGSMSEYEDIRVPFVVRGPNTEAG---CLNPHVILNIDLAPTILDIAG-----374
Db 332 HRCAGKVTAYDEDLRVFPLIRGGRASHSDKPNASKVLHVDPAFTILTLAGAGDQVG 391
2y 375 -----LDIPADMKGSKILKLDTERPVRNPHLKKRVRDSEFLVE-----RGLKLH 421
Db 392 DKALDGTPLGLYANDDGN---LLADYVRPANH-----RNQFGFWGWSDEVILH 438
2y 422 KRDNKVDAAEENFLPKYQ-----RVKDLQCORAEVQACEQKGKQWCEVADATGKLKL 474
Db 439 -----HIPRYTNSWKAVRYD-----EDNQAWKLIYVCTNEREL 474
2y 475 HKCK 478
Db 475 YDLK 478

RESULT 4
ARS-CHLRE STANDARD; PRT; 647 AA.
AC P14217;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
GN AS.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-55.
RC STRAIN=cw15;
RX MEDLINE=9384447; PubMed=2476654;
RA de Hostos E.L., Schilling J., Grossman A.R.;
RT "Structure and expression of the gene encoding the periplasmic
EL arylsulfatase of Chlamydomonas reinhardtii.";
RN Mol. Gen. Genet. 218:229-239(1989).
RP [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFTS.
RA Bauroch A.;
RL Unpublished observations (AUG-1996).
CC -!- CATALYTIC ACTIVITY: A phenol sulfite + H(2)O = a phenol + sulfate.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- INDUCTION. By sulfur deprivation.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. FRAMESHIFTS HAD TO BE
CC INTRODUCED TO MAXIMIZE THE SIMILARITY WITH OTHER SULFATASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X16180; CAA34302.1; ALT_FRAME.
```

```
DR EMBL; X16179; CAA34301.1; ALT_FRAME.
DR EMBL; X52304; CAA36545.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 21
FT CHAIN 22 647
FT MOD_RES 73 73
FT CARBOHYD 42 42
FT CARBOHYD 90 90
FT CARBOHYD 279 279
FT CARBOHYD 444 444
FT CARBOHYD 528 528
SQ SEQUENCE 647 AA; 72102 MW; 7404EAB1EF233F0B1 CRC64;
Query Match 8.5%; Score 405; DB 1; Length 647;
Best Local Similarity 26.7%; Pred. No. 1.3e-21;
Matches 132; Conservative 76; Mismatches 179; Indels 108; Gaps 20;
QY 43 RPNILVLTDQDVELGSM--QVMNKTTRIMEQGAFFINAFVTTMCCPSRSILTKY 100
Db 26 KPNFWLFTDDQDAIQNSTHPIYMPSLHKYIRYPGVLSQYFVTTVPCCPSRTNLARGQF 85
QY 101 VHNHTVTNNENCSPP-----SWQAQHESETP-AVILNSTGYETAFFGKYLNEY--NGS 151
Db 86 AHN-----TNFTSVLPYGGWAKWKGIDQSYLPLMLKQGYNTYVVGKFLVDYSVNYQ 141
QY 152 YVP-PGWKEWVGLLKNSRFNYTLCRNGVKRSGSDYSKDYLTDLITNDSVSPRTSKM 210
Db 142 QVPRAGTISMEXVTPYTFDYNTRLQRNGATP---NIYPGEYSTDVIRDKGVAQIKSA--V 196
QY 211 YPHRPVLMTSHAAPGPDSPAQYSLFPNASQHTPTPSYNYAPNP-DKHWIMRYTGP- 268
Db 197 AAGKFFYAQISPIAPH-----TSTQISTNPATGVTSTRSYFFPIAPPWHQLFSDANLP 249
QY 269 -----KEI-----HMEFTNMLQRKRLQTLMSVDDSMETIYNMLV 302
Db 250 GGSXKNLYEVDSKPAWIRALPLAQNNRTQBEIYRLRLSL-GPDELIQVVKTL 308
QY 303 ETGELDNITYVYTDHGHVHGGFLVKGSMPEFDIRVPFVVRGPNVEAG-----CLNPH 358
Db 309 EAGVLDNTYIIYSADNGYHVGARFGAGKTGYEEDLRVFLIRGPGIKASKSDXPNSK 368
QY 359 IVLNIDLAPTILDIAG-----LDIPADMKGSKILKLDTERPVRNPHLKKX 404
Db 369 VGLHVDPAFTILTLAGASHILGDKLDTGLGLYANDDG-----TLPSDYPRPEHQHQF 423
QY 405 MRV-----WRDSFLVERGKLLHKRDNKVDQAENFLPKYQVXKDLQORASYQTACEQLGQ 460
Db 424 QGEFWGGSDELL-----QNLRSQPNNTW-KVVRTYD-----ESSKQ 459
QY 461 KWQCVEDATGKLKLH 475
Db 460 GWKLIACQCTNERELY 474

RESULT 5
YIDJ_ECOLI
ID YIDJ_ECOLI STANDARD; PRT; 497 AA.
AC P31447;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative sulfatase yidJ (EC 3.1.1.6.-).
GN YIDJ OR B3678.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
```



158 --ITNOMEYDDEVAFLANQKLYLSRENDESRFPWCLTVSFTBHPDPVARRKFWDLIYE 215  
 241 NASQHTPSYNYAP-----NPDKXHWIMRYTGPMKPIHMFETNM-----LQKRB---LQTL 287  
 216 DC-EHUTPEVGAIFLDEQPHSQRM-----JSCDYQNFVDTEENVRRSRRAYFANI 266  
 288 MSVDDSMETIYNMIVETGELDNITYVYTAGHGYHIGQFGLVKGKSMPEYFDIRVPFYVRG 347  
 267 SYLDEKVGELIDTLNTRMLDDTLILFCSHDGMDLGERGL-WFKMNFEGSGARVPLMIAG 325  
 348 PNVFACG-INPHVILNIDAPILDIAGL---DIPADMKGKILKLLD 391  
 326 PGIAPGLHLP--TSNLDVTPTLADLAGISLEVRPWTGDVSLVPMVN 371

## RESULT 7

16-OCT-2001 (Rel. 40, Created)  
 16-OCT-2001 (Rel. 40, Last sequence update)  
 16-OCT-2001 (Rel. 40, Last annotation update)  
 Putative sulfatase yden precursor (EC 3.1.6.-).  
 YDEN OR B1498.  
 3N Escherichia coli.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Escherichia.  
 CC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12; MG1655;  
 RC MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RA "The complete genome sequence of Escherichia coli K-12.";  
 RT Science 277:1459-1474 (1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RC MEDLINE=97251357; PubMed=9097039;  
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Moromura K.,  
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,  
 RA Sampaio G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RT DNA Res. 3:363-377 (1996).  
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE000247; AAC74571.1; ALT INIT.  
 DR EMBL; D90791; BAA15169.1; ALT INIT.  
 DR EMBL; D90792; BAA15172.1; ALT\_INIT.  
 DR EcoGene; EG13796; yden  
 DR InterPro; IPR000917; Sulfatase.  
 DR Pfam; PF00884; Sulfatase; 1.  
 DR PROSITE; PS00523; SULFATASE\_1; FALSE NEG.  
 DR PROSITE; PS00149; SULFATASE\_2; 1.  
 KW Hypothetical protein; Hydrolase; Signal; Complete proteome.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT CHAIN 30 560 PUTATIVE SULFATASE YDEN.  
 FT ACT\_SITE 185 185 POTENTIAL.

SQ SEQUENCE 560 AA; 62802 MW; 67ED7FF7696C7A9F CRC64;  
 Query Match 5.3%; Score 251; DB 1; Length 560;  
 Best Local Similarity 22.7%; Pred. No. 1.5e-10;  
 Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;  
 QY 18 LCGSSAFISHHLKRGFORDRNI-----RPMILVLTDDQ----- 54  
 DB 18 LASGMAFAAAHADDVKLKATKTNVAFSDFTETEXTKGNPIIIVLTMDDLGYGQLPFDK 77  
 QY 55 -----DVELGSMVMKTRIMEQGAHFNAFVTTPMCSPRS 93  
 DB 78 GSDFPKTMENREVDTYKIGDKAIEAAQKSPITLLSINDEGVRFNGVYAHGVSGPSRA 137  
 QY 94 SILTKYVHNHNTYNNENCSSPSWOAHE---SRTFAVYL-NSTGYRTAFPGK-VILNEY 148  
 DB 138 AINTGRAPARFGVYSNTD-----AQDGIPLTETFLPQLFQNHGYTAAVKGWHLSKI 189  
 QY 149 NGSVYPPGKEWGLLKNSRFY--NYTL--CRNGVKEHGSY----- 187  
 DB 190 SNVPVPED-----KQTRDYHDNFTTSAEHWQPNRGFDYFMGFHAGTAYNPSL 241  
 QY 188 -----SKDYLTDLITNDSVSFFRTSKMYPHREVLVMSHAAPHGPEDS-AP-QYSR 237  
 DB 242 FKNERVPAKGIYSQDLTDEAIGVDRAKTL--DQFFMLYLAYNAPHLPNENAPDQYOK 299  
 QY 238 LFPNASHIPTPSYNYAPNPKHWMRYTGMKPIEHMEFTMLQKELQILMSVDDSMETI 297  
 DB 300 QFNTGSQ--TADNYA-----SVYSVDQGVKRI 325  
 QY 298 YNMLVETGELDNITYVVTADHGYHI--GQF---GLVKG-KSMPEYFDIRVPFYVRGPNVEA 352  
 DB 326 LEQLKNGQVDTNIIILFTSDNGAVIDGLPLNGAQKYSQTYPGGTHTFMFMW---K 381  
 QY 353 GCLNP-----HIVNIDILAPITLIDAGLIDPAD--MDGKSILKLLDTER---PVNRPHLKK 403  
 DB 382 GKLPQNGYDKLISAMDYPTALDAADISIPKDLKLDGVSLLPWLDKQKQEPHKNITWIT 441  
 QY 404 KMRVWRDSFLVERGKLHK-----RNDN--KVDAQENFL 436  
 DB 442 SYSHWFDEENIPFDNHYKFRHQSDDYPHNPNTEDLSQPSYTVRNNDYSLVTVNNQL 501  
 QY 437 PKYQVRKDLQRAEYQFACELQKQKQCV 465  
 DB 502 GLY-KLTDLQCKNLAAANPQVVKEMQGV 529  
 RESULT 8  
 STS\_HUMAN STANDARD; PRT; 583 AA.  
 ID STS\_HUMAN AC P08842;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfatase) (Steryl-  
 GN sulfate sulfohydrolase) (Arylsulfatase C) (ASC).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89340479; PubMed=2668275;  
 RA Stein C., Hille A., Seidel J., Rijnbout S., Waheed A., Schmidt B.,  
 RA Geuze H., von Figura K.;  
 RT "Cloning and expression of human steroid-sulfatase. Membrane  
 RT topology, glycosylation, and subcellular distribution in BHK-21  
 RT cells.";  
 RL J. Biol. Chem. 264:13865-13872 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87187642; PubMed=3032454;



Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T., Shapiro L.J.;  
"Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y interchange."; Cell 49:443-454(1987).  
(3)  
SEQUENCE OF 134-174 AND 461-583 FROM N.A.  
MEDLINE=89077541; PubMed=3203382;  
Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L., Neiswanger K., Shapiro L.J.;  
"The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution."; Cell 55:1123-1135(1988).  
(4)  
SEQUENCE OF 22-45.  
TISSUE=Liver;  
MEDLINE=89352671; PubMed=2765556;  
Kawano J.-I., Kotani T., Ohtaki S., Minamino N., Matsuo H., Oinuma T., Aikawa E.;  
"Characterization of rat and human steroid sulfatases."; Biochim. Biophys. Acta 997:199-205(1989).  
(5)  
VARIANTS XLI LEU-341; ARG-372 AND TYR-446.  
MEDLINE=92170784; PubMed=1539590;  
Basler E., Grompe M., Parenti G., Yates J., Ballabio A.;  
"Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis."; Am. J. Hum. Genet. 50:483-491(1992).  
(6)  
VARIANTS XLI LEU-341; ARG-372; SER-372; ARG-444 AND TYR-446.  
MEDLINE=97400563; PubMed=9252398;  
Alperin E.S., Shapiro L.J.;  
"Characterization of point mutations in patients with X-linked ichthyosis. Effects on the structure and function of the steroid sulfatase protein."; J. Biol. Chem. 272:20756-20763(1997).  
(7)  
VARIANT XLI PRO-560.  
MEDLINE=20146224; PubMed=10679952;  
Sugawara T., Shimizu H., Hoshi N., Fujimoto Y., Nakajima A., Fujimoto S.;  
"PCR diagnosis of X-linked ichthyosis: identification of a novel mutation (S560P) of the steroid sulfatase gene."; Hum. Mutat. 15:296-296(2000).  
(8)  
VARIANT XLI ARG-380.  
MEDLINE=20304877; PubMed=10844566;  
Oyama N., Satoh M., Iwatsuki K., Kaneko F.;  
"Novel point mutations in the steroid sulfatase gene in patients with X-linked ichthyosis: transfection analysis using the mutated genes."; J. Invest. Dermatol. 114:1195-1199(2000).  
-!- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO ESTROGENS DURING PREGNANCY.  
-!- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate + H<sub>2</sub>O = 3-beta-hydroxyandrost-5-en-17-one + sulfate.  
-!- SUBUNIT: Homodimer.  
-!- SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE. THE SEQUENCE SHOWS SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE PROTEIN IN THE MICROSOMAL MEMBRANE.  
-!- DISEASE: Defects in STS are the cause of x-linked ichthyosis (XLI), a diskretization disorder characterized by the presence of prominent scales, especially on the neck, extremities, trunk, and buttocks. It affects approximately 1 of 2000-6000 males.  
-!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
CC EMBL; J04964; AAA60597.1; -  
CC EMBL; M16505; AAA60596.1; -  
DR EMBL; M23945; AAA60598.1; -  
DR EMBL; M23567; AAA60599.1; -  
DR PIR; A32641; KJHUAC.  
DR HSP; P15848; IFSU.  
DR Genew; HGNC:11425; STS.  
DR MIM; 308100; -  
DR GO; GO:0005783; C:cytoplasmic reticulum; TAS.  
DR GO; GO:0005768; C:cytosome; TAS.  
DR GO; GO:0005794; C:Golgi apparatus; TAS.  
DR GO; GO:0005764; C:lysosome; TAS.  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0005792; C:mitochondrion; TAS.  
DR GO; GO:0005886; C:plasma membrane; TAS.  
DR GO; GO:0004773; P:steroid sulfatase activity; TAS.  
DR GO; GO:0008544; P:epidermal differentiation; TAS.  
DR GO; GO:0006706; P:steroid catabolism; TAS.  
DR InterPro; IPR000917; Sulfatase.  
DR Pfam; PF00884; Sulfatase; 1.  
DR PROSITE; PS00523; SULFATASE\_1; 1.  
DR PROSITE; PS00149; SULFATASE\_2; 1.  
KW Hydrolyase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;  
KW Pregnancy; Signal; Disease mutation.  
FT SIGNAL 1 21  
FT CHAIN 22 583  
FT MOD\_RES 75 75  
FT STERYL-SULFATASE.  
FT 2-AMINO-3-OXOPROPIONIC ACID (BY  
FT SIMILARITY).  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT S -> L (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007240.  
FT W -> R (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007241.  
FT W -> S (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 014020.  
FT G -> R (IN XLI).  
FT /FTID=VAR 014021.  
FT H -> R (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 014022.  
FT C -> Y (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007242.  
FT Q -> P (IN XLI).  
FT /FTID=VAR 014023.  
FT A -> E (IN REF. 2).  
FT CONFLICT 23 23  
SQ SEQUENCE 583 AA; 65492 MW; 74746AFA9D21A0A6 CRC64;  
Query Match 5.0%; Score 238.5; DB 1; Length 583;  
Best Local Similarity 20.7%; Pred. No. 1.3e-09;  
Matches 126; Conservative 67; Mismatches 180; Indels 237; Gaps 23;  
QY 43 RPNILAVLTDODVGLSGVMV-NKTRRI-----MEQGAHAFVVTTPMCCPSRSSIL 96  
Db RPNILVLMAD-DLIGDPCGVNKTIRTPNDRLASGVGVKUTQHLAASPLCTPSRAAFM 83  
QY 97 TGYVENVHTYNNENCSFSPWQ-----ACHESRTFAVYLNSTGYRTAFQK 143  
Db TGRY-----PVRSGMASWSRTGVFLFTASSGGLPTDEITPAKLLKDGQSTALIGK 134  
QY 144 Y-----INEY-----NGSYVPPGNKE----- 159  
Db WHLGMSCHSKTDFCHHPLHHGFNYFVIGISLTNLRCKPGEGSVFTTGFRLVFLQIVG 194  
QY 160 -----WVGLLKNSRFVNYTLCRNGVKRKHG 184  
Db VTLLTLAALNCIGLLHVLGVFSLFLAALTLTLFLGLFHYFRPLNCFMKN--YEIIQ 252

```

QY 185 SDYSKDYLTDLITNDSVSPERTSKQYPHRPVLMVISHAAPHGPDSPAPQYSLRPENASQ 244
DB 253 QPMSYNDLQRLTVEAAQFIQRNTE-----TFPLVLVLSLVH-----TALFSSKDFAGSQ 304
QY 245 HITPSYNAAPNPKHIMRYTGMKPIHMEFTNMLQKLOTLMSVDDSMETIYNMLVET 304
DB 305 H--GVYGDA-----VEEMDSVGGQILNLLDEL 329
QY 305 GELDNTIVVTADGHYI-----QQFGLVK-GKSMPEFDPVFPVYVVGPNV-E 351
DB 330 RLANDTLIYTSQGAHVEVSKGEIHGSGNIYKGGKANNWEGGIRVPGILRPVVIQ 389
QY 352 AGCLNPHVINIDILAPLIDIPAD--MDGKSILKLD--TERPVN----- 397
DB 390 AGQKIDEPSTNMDIFPTVAKLAGAPLPEDRIIDGRDLMPLEKQSRSDHEFLFYHCNAY 449
QY 398 -----REFLKCKMRVWRDSFLVER-----GKLLHRD-----NDKVD 430
DB 450 LNAVREHPQNSTIWKAFPTFPNPNVPGSNGCPATHVCFGSGYVTHDPPLLFDISKDP 509
QY 431 QEENFL-----PKYQVKDLCQRAE-----YOTACEOLGQK 461
DB 510 RERNPLTFAPEPRYEYLLKVMQEADEHTQTLPEVPDQFSWNNFLKMWLQCCPSTGLS 569
QY 462 WQCVEDATGK 471
DB 570 CQCDREKQDK 579

RESULT 9
ARS_STRPU STANDARD; PRT; 567 AA.
ID ARS_STRPU
AC PS0473;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase) (ARS).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
CX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89357267; PubMed=2767335;
RA Yang Q., Angerer L.M., Angerer R.C.;
RT "Structure and tissue-specific developmental expression of a sea urchin arylsulfatase gene.";
RL Dev. Biol. 135:53-65(1989).
CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSION IS CONFINED TO ABORAL ECTODERM CELLS AND THEIR PRECURSORS.
CC -!- DEVELOPMENTAL STAGE: LOW LEVELS ARE FOUND AT MESENCHYME BLASTULA STAGE (24 HR), LEVELS INCREASE BY LATE GASTRULA STAGE AND ARE MAINTAINED AT PLUTEUS STAGE.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28404; AAA30036.1; -
DR PIR; A37362; A37362.
DR HSP; F15289; 1AUK.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.

```

```

DR PROSITE; PS00523; SULFATASE 1; 1.
KW Hydrolase; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 567 ARYLSULFATASE.
FT MOD_RES 115 115 2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
FT ACT_SITE 173 173 POTENTIAL.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 567 AA; 62477 MW; D55B6279834C4D6 CRC64;

Query Match 5.0%; Score 238; DB 1; Length 567;
Best Local Similarity 23.4%; Pred. No. 1.4e-09;
Matches 103; Conservative 58; Mismatches 167; Indels 112; Gaps 18;

QY 2 GPPSLVLCILSATVFSLLGGSSAFLSHRLKGRFQDRNRIRNIIILVLTDDQDVLGSM 61
DB 27 GPDAESLASLDRTATRYGDGDL--HLLGQGGHRTAMTKENVILLADDDGV--GDL 82
QY 62 QVMNKTR-----IMEQGAHFNAFVTTMCPSSRSILTKYVHNHNTYTNN----- 111
DB 83 SVYGHPTCEPGFIDQMANQGLRFTQGYSGDSVCTPSRSAIVTGRQPIRTGVYGEERFLP 142
QY 112 --NCSSPSWQAQHSRTFAVILNSTGYRTAFQKY-----LNE-----YNGSYVP----- 154
DB 143 WTTTGLPLYEV-----TIAEMKAGAGYTTGVGKWHLGINENSSDGAHLPAKRGPFVVG 197
QY 155 --PGWKEW-----VGLLKN-----SRFNYTLCRNGVKEKHGSDYSKDYLTDLITNDS 200
DB 198 HNLFPGNWRCDDTGLHQDPDPTNACFLYNSVTAQPFQHG-----LTQLLRD 249
QY 201 VSFFRTSKMYPHPVLMVISHAAPHGPDSPAPQYSLFPNASQHIPTPSYNAAPNPKHW 260
DB 250 VGFIEDN-----VKNPFPMYVFAHMH-----TSLFSSD----- 278
QY 261 IMRYTGMKPIHMEFTNMLQKRL-QTLMVDDSMETIYNMLVETGBELDNTIYVVTADHG 319
DB 279 -----DFSCSTRGRYGDNLREMDQAEIQIVTLVDNDIDDDNTVIFFTSDHG 325
QY 320 YH---IGQFGLVK-----GKSMPEFDPVFPVYVVGPNVEAGCLNPHVILNIDLAPTLDI 372
DB 326 PHREYCGEGGDANVFRGGKQSGMEGGRHPYIVYVWPGTISPGVSHSHEIVTSMIDIATAVNL 385
QY 373 AGLDIPAD--MDGKSILKLD 390
DB 386 GGSOLPTDRIYDGKLSVL 405

RESULT 10
IDS_HUMAN
ID IDS_HUMAN STANDARD; PRT; 550 AA.
AC P22304; Q14604;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iduronate 2-sulfatase precursor (EC 3.1.6.13).
GN IDS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-58 AND 456-473.
RX MEDLINE=91046030; PubMed=2122463;
RA Wilson P.J., Morris C.P., Anson D.S., Ochiodoro T., Bielicki J.,
RA Clements P.R., Hopwood J.J.;
RT "Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8531-8535(1990).
RN [2]

```

- IP SEQUENCE FROM N.A.  
UX MEDLINE=94063929; PubMed=8244397;  
LA Wilson P.J., Meaney C.A., Hopwood J.J., Morris C.P.;  
LA "Sequence of the human iduronate 2-sulfatase (IDS) gene";  
IL Genomics 17:773-775(1993).  
IL [3]  
IP SEQUENCE FROM N.A.  
P MEDLINE=96352905; PubMed=8717057;  
X Timms K.M., Lu F., Shen Y., Pierson C.A., Muzny D.M., Gu Y.,  
A Nelson D.L., Gibbs R.A.;  
A "130 kb of DNA sequence reveals two new genes and a regional  
T duplication distal to the human iduronate-2-sulfatase  
T locus";  
T Genome Res. 5:71-78(1995).  
L [4]  
IP SEQUENCE FROM N.A.  
P Muzny D., Ansari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,  
A Shen Y., Rowland K., Liu W., Perez L., Ding Y., Gonzalez O.,  
A Haywood M., Jain A., Leal B., Logan O., Nguyen V., Savage L., Shen H.,  
A Worley K., Chen E., Forcum J., Aronson A.D., Chiu M.W., Gorrell J.H.,  
A Brundage E., Di W., Chinalic C., Nelson D., Gibbs R.A.;  
L Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
IL [5]  
IP SEQUENCE FROM N.A. (ISOFORM SHORT).  
P TISSUE=Lymphocytes;  
C MEDLINE=96079126; PubMed=8530090;  
C Malmgren H., Carlberg B.M., Pettersson U., Bondeson M.L.;  
A "Identification of an alternative transcript from the human iduronate-  
T 2-sulfatase (IDS) gene";  
T Genomics 29:291-293(1995).  
L [6]  
IP SEQUENCE OF 1-398 FROM N.A.  
P MEDLINE=93258349; PubMed=8490623;  
C Flomen R.H., Green E.P., Green P.M., Bentley D.R., Giannelli F.;  
A "Determination of the organisation of coding sequences within the  
T iduronate sulphate sulphatase (IDS) gene";  
T Hum. Mol. Genet. 2:5-10(1993).  
L [7]  
IP REVIEW ON MPS-II VARIANTS.  
P MEDLINE=94154729; PubMed=8111411;  
X Hopwood J.J., Bunge S., Morris C.P., Wilson P.J., Steglich C.,  
A Beck M., Schwinger E., Gal A.;  
A "Molecular basis of mucopolysaccharidosis type II: mutations in the  
T iduronate-2-sulphatase gene";  
T Hum. Mutat. 2:435-442(1993).  
L [8]  
IP VARIANTS MPS-II ARG-135 AND GLY-422.  
X MEDLINE=93265059; PubMed=1303211;  
X Bunge S., Steglich C., Beck M., Rosenkranz W., Schwinger E.,  
A Hopwood J.J., Gal A.;  
A "Mutation analysis of the iduronate-2-sulfatase gene in patients with  
T mucopolysaccharidosis type II (Hunter syndrome).";  
T Hum. Mol. Genet. 1:335-339(1992).  
L [9]  
IP VARIANT MPS-II TRP-468.  
P MEDLINE=93258320; PubMed=1284597;  
X Crotti P.L., Bunge S., Anderson R.A., Whitley C.B.;  
A "Mutation R468W of the iduronate-2-sulfatase gene in mild Hunter  
T syndrome (mucopolysaccharidosis type II) confirmed by in vitro  
T mutagenesis and expression";  
T Hum. Mol. Genet. 1:755-757(1992).  
L [10]  
IP VARIANTS MPS-II ARG-86; ASP-94; ARG-120; PRO-221 AND GLY-422.  
X MEDLINE=94108441; PubMed=8281149;  
X Bunge S., Steglich C., Zuther C., Beck M., Morris C.P., Schwinger E.,  
A Schinzel A., Hopwood J.J., Gal A.;  
A "Iduronate-2-sulfatase gene mutations in 16 patients with  
T mucopolysaccharidosis type II (Hunter syndrome).";  
T Hum. Mol. Genet. 2:1871-1875(1993).  
L [11]  
IP VARIANTS MPS-II GLU-68; HIS-293; GLY-478 AND ARG-485.  
X MEDLINE=95072615; PubMed=7981716;  
X Schroeder W., Wulff K., Wehnert M., Seidlitz G., Herrmann F.H.;  
A
- "Mutations of the iduronate-2-sulfatase (IDS) gene in patients with  
Hunter syndrome (mucopolysaccharidosis II).";  
Hum. Mutat. 4:128-131(1994).  
[12]  
VARIANTS MPS-II TRP-132; TYR-229; ARG-358; HIS-469 AND CYS-523.  
RX MEDLINE=95193786; PubMed=7887413;  
RA Jonsson J.J., Aronovich E.L., Braun S.E., Whitley C.B.;  
RT "Molecular diagnosis of mucopolysaccharidosis type II (Hunter  
RT syndrome) by automated sequencing and computer-assisted  
RT interpretation: toward mutation mapping of the iduronate-2-sulfatase  
RT gene";  
RT Am. J. Hum. Genet. 56:597-607(1995).  
RN [13]  
VARIANTS MPS-II LEU-86; ASN-87; PRO-92; ASN-135; CYS-345 AND TRP-468.  
RX MEDLINE=95245347; PubMed=7728156;  
RA Popowska E., Rathmann M., Tytki-Szymanska A., Bunge S., Steglich C.,  
RA Schwinger E., Gal A.;  
RT "Mutations of the iduronate-2-sulfatase gene in 12 Polish patients  
RT with mucopolysaccharidosis type II (Hunter syndrome).";  
RL Hum. Mutat. 5:97-100(1995).  
RN [14]  
VARIANT MPS-II VAL-346.  
RX MEDLINE=95322987; PubMed=7599640;  
RA Li P., Huffman P., Thompson J.N.;  
RT "Mutations of the iduronate-2-sulfatase gene on a T146T background in  
RT three patients with Hunter syndrome";  
RL Hum. Mutat. 5:272-274(1995).  
RN [15]  
VARIANTS MPS-II.  
RX MEDLINE=97094177; PubMed=8940265;  
RA Rathmann M., Bunge S., Beck M., Kresse H., Tytki-Szymanska A., Gal A.;  
RT "Mucopolysaccharidosis type II (Hunter syndrome): mutation 'hot spots'  
RT in the iduronate-2-sulfatase gene";  
RL Am. J. Hum. Genet. 59:1202-1209(1996).  
RN [16]  
VARIANTS MPS-II LEU-333 AND ASP-346.  
RX MEDLINE=96163494; PubMed=8566953;  
RA Olsen T.C., Eiken H.G., Knappskog P.M., Kase B.F., Mansson J.-E.,  
RA Boman H., Apold J.;  
RT "Mutations in the iduronate-2-sulfatase gene in five Norwegians with  
RT Hunter syndrome";  
RL Hum. Genet. 97:198-203(1996).  
RN [17]  
VARIANTS MPS-II LEU-333 AND GLY-334.  
RX MEDLINE=96255081; PubMed=8830188;  
RA Li P., Thompson J.N.;  
RT "Detection of four novel mutations in the iduronate-2-sulphatase gene  
RT by single-strand conformation polymorphism analysis of genomic  
RT amplicons";  
RL J. Inherit. Metab. Dis. 19:93-94(1996).  
RN [18]  
VARIANTS MPS-II ASP-63; THR-347; GLN-468 AND LEU-468.  
RX MEDLINE=97365936; PubMed=9222763;  
RA Vallani G.R.D., Balzano N., Grosso M., Salvatore F., Izzo P.,  
RA di Natale P.;  
RT "Mucopolysaccharidosis type II: identification of six novel mutations  
RT in Italian patients";  
RL Hum. Mutat. 10:71-75(1997).  
RN [19]  
VARIANT MPS-II GLN-468.  
RX MEDLINE=98041699; PubMed=9375851;  
RA Sukegawa K., Song X.-Q., Masuno M., Fukao T., Shimozawa N., Fukuda S.,  
RA Isogai K., Nishio H., Matsuo M., Tomatsu S., Kondo N., Orii T.;  
RT "Hunter disease in a girl caused by R468Q mutation in the  
RT iduronate-2-sulfatase gene and skewed inactivation of the X  
RT chromosome carrying the normal allele";  
RL Hum. Mutat. 10:361-367(1997).  
RN [20]  
VARIANTS MPS-II ASN-45; TYR-115; LEU-228; ARG-266; LYS-434; LYS-485  
RX AND CYS-502.  
RX MEDLINE=99092178; PubMed=9875019;  
RA Vafiadaki E., Cooper A., Heptinstall L.E., Hatton C.E., Thornley M.,  
RA Wraith J.E.;

"Mutation analysis in 57 unrelated patients with MPS II.";  
Arch. Dis. Child. 79:237-241(1998).

[21]

VARIANTS MPS-II.

MEDLINE=99118896; PubMed=9921913;

Karssten S., Voskoboova E., Tishkanina S., Petersson U.,

Krasnopolskaja X., Bondeson M.-L.;

"Mutational spectrum of the iduronate-2-sulfatase (IDS) gene in 36

unrelated Russian MPS II patients.";

Hum. Genet. 103:732-735(1998).

[22]

VARIANTS MPS-II LEU-86; HIS-88; PRO-88; ILE-118 AND HIS-266.

Balzano N., Villani G.R.D., Grosso M., Izzo P., di Natale P.;

"Detection of four novel mutations in the iduronate-2-sulfatase

gene.";

Hum. Mutat. 11:333-333(1998).

[23]

VARIANTS MPS-II THR-85; HIS-88; ILE-349 AND VAL-521.

MEDLINE=98112423; PubMed=9452044;

Gott L., Coll M.J., Chabas A.;

"Mutations in the iduronate-2-sulfatase gene in 12 Spanish patients

with Hunter disease.";

Hum. Mutat. Suppl. 1:866-S68(1998).

[24]

VARIANTS MPS-II PHE-143; TRP-184; VAL-269 AND HIS-348.

Karssten S.L., Voskoboova E., Carlberg B.-M., Kleijer W.J.,

Toennesen T., Petersson U., Bondeson M.-L.;

"Identification of 9 novel gene mutations in 19 unrelated Hunter

syndrome (Mucopolysaccharidosis type II) patients.";

Hum. Mutat. 12:433-433(1998).

[25]

VARIANTS MPS-II PHE-143; LYS-341; TYR-342 AND PHE-491.

MEDLINE=99235558; PubMed=10220152;

Query Match 4.9%; Score 234.5; DB 1; Length 550;

Best Local Similarity 21.2%; Pred. No. 2.3e-09;

Matches 114; Conservative 80; Mismatches 189; Indels 155; Gaps 23;

2y 1 MGPPS-----LVLLLSATVFLSGSSAFLSHRLKGRFDRNRNITLVLTDDQ 55

2b 1 MPPTTGRLLWGLVSSVCVALGSET-----QANSTDAINLVIIIVDLR 48

2y 56 VELGSMQVKNKTRRIMEQGAH---FINAVTTPMCCPSRSILTKG-----YVHNHTY 107

2b 49 PSLGCGDKLVRSFNIDQLASHLLFONAFQAQVCAPEKSVFLGRRDPTTLRYDFNSY 108

2y 108 TNNENCSPPSWQAQHS-RTFAVILNSTGYRTAFFGKYL-----NEYNGSYVP- 154

2b 109 -----WRVHAGNFSTIPQYFKENGVTMSVGKVFHPGICSSNHTDDSPYSWFFPY 158

2y 155 -PGWKEMVGLLKNSRPNVNTLCRNGVKEKHGS-----DYSKDYLTD-LITNDSVGF 204

2b 159 HP-----SSEKENTKCEGPGELHANLLCEVDVLDVPEGLTPKQSTEQAIQLL 209

2y 205 RTGKQVHPRVLMVTSAAHPGPDSAPQYGRFLFNASQHTTPSNVAPND----- 257

2b 210 EKQMT--SASPPFLVGVYKHPHFIPRYPKFOKLYP--LENIT----LAPDPEVDGLPP 261

2y 258 ---KHWI-----MRVTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNVLVET 304

2b 262 VAYNPMDIQRQEDVQALNISVPGYIPVDFQKIRQSYFASVYLDVTQVGRLLSALDDL 321

2y 305 GELDNTYIVVTDAGHYHIGQGLVKGSKMPEYFDIRVP--FVYRGN----- 349

2b 322 QLANSTIIAFTSDHGWALGEBG--EWAKYENFNVATHVPLIFVPGRTASLPEAGEKLPFY 380

2y 350 -----VEAGCLNPHVLNIDLAPILDTAGLDDIADMDGKSLKLLTETPVNRP 399

2b 381 LDPDFSASQMLPEGROSMDLVELVSLFPLTAGLQVLP-----PRCFVPSF 427

2y 400 HLKKQVWRDPSFLVRGKLLHRRKNDKVDQAENFLP-----KYQVRKDLQ 447

2b 428 HVE-----LCREGKLLKHFRF--DLEDPPVLPFGNPRELAYSQYPPSPDIQ 474

RESULT 11

ARS\_HEMPU

ID ARS\_HEMPU STANDARD; PRT; 551 AA.

AC P14000;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphonylhydrolase)

(ARS).

OS Hemocentrotus pulcherrimus (Sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidea; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;

OC Hemocentrotus.

OX NCBI\_TaxID=7650;

RN [1]

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Pluteus;

RX MEDLINE=89030699; PubMed=3181160;

RA Sasaki H., Yamada K., Akasaka H., Suzuki K., Saito A., Sato M.,

Shimada H.;

"cDNA cloning, nucleotide sequence and expression of the gene for

arylsulfatase in the sea urchin (Hemocentrotus pulcherrimus)

embryo.";

RL Eur. J. Biochem. 177:9-13(1988).

RN [2]

SEQUENCE FROM N.A.

RX MEDLINE=90092130; PubMed=2598936;

RA Yamada K., Akasaka K., Shimada H.;

"Structure of sea-urchin arylsulfatase gene.";

RL Eur. J. Biochem. 186:405-410(1989).

CC -1- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE EXTRACELLULAR

MATRICES.

CC -1- MATRICES INVOLVED IN CELL MOVEMENT DURING MORPHOGENESIS.

CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.

CC -1- SUBCELLULAR LOCATION: IN BOTH THE CYTOPLASM AND THE EXTRACELLULAR

MATRICES.

CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X17015; CAA34881.1; -

DR PIR; S01793; S01793.

DR HSPP; P15289; IAUU.

DR InterPro; IPR000917; Sulfatase.

DR Pfam; PF00884; Sulfatase; 1.

DR PROSITE; PS00523; SULFATASE\_1; 1.

DR PROSITE; PS00149; SULFATASE\_2; 1.

KW Hydrolase; Signal; Glycoprotein; Extracellular matrix.

FT SIGNAL 1 20

FT CHAIN 21 551

FT MOD\_RES 21 21

FT MOD\_RES 100 100

FT ACT\_SITE 158 158

FT CARBOHYD 164 164

FT CARBOHYD 213 213

FT CARBOHYD 296 296

FT SEQUENCE 551 AA; 60952 MW; 54C1AAC14D6710C9 CRC64;

Query Match 4.9%; Score 232; DB 1; Length 551;

Best Local Similarity 23.8%; Pred. No. 3.5e-09;

Matches 105; Conservative 63; Mismatches 161; Indels 116; Gaps 21;

Qy 42 IRPNIILVLTDDQDVELGSMQVKNKTRRIMEQGAHFINAFVTPMCCPSRS 93

Db 50 VKENVLLVAD-----HNGSGDLTSYGHPTQEAQFIDKMAEGLRFTNGYVGDVCTPSRS 105



RA Hori T., Iwata H., Ogawa T., Nakashima Y., Hanyu Y., Hashimoto T.,  
RA Titani K., Oyama R., Suzuki M., Yagi K., Hayashi Y., Orii T.,  
RT "Morquio disease: isolation, characterization and expression of full-  
RT length cDNA for human N-acetylgalactosamine-6-sulfate sulfatase.";  
RN Biochem. Biophys. Res. Commun. 181:677-683(1991).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95095267; PubMed=8001980;  
RA Morris C.P., Guo X.H., Apostolou S., Hopwood J.J., Scott H.S.;  
RT "Morquio A syndrome: cloning, sequence, and structure of the human N-  
RT acetylgalactosamine 6-sulfatase (GALNS) gene.";  
RN Genomics 22:652-654(1994).  
[3]  
RP VARIANT MPS-IVA LYS-204.  
RX MEDLINE=92395122; PubMed=1522213;  
RA Fukuda S., Tomatsu S., Masue M., Sukegawa K., Iwata H., Ogawa T.,  
RA Nakashima Y., Hori T., Yamagishi A., Hanyu Y., Morooka K., Kitan T.,  
RA Hashimoto T., Orii T.;  
RT "Mucopolysaccharidosis type IVA. N-acetylgalactosamine-6-sulfate  
RT sulfatase exonic point mutations in classical Morquio and mild  
RT cases.";  
RL J. Clin. Invest. 90:1049-1053(1992).  
[4]  
RP VARIANTS MPS-IVA.  
RX MEDLINE=95397840; PubMed=7668283;  
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,  
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,  
RA Suzuki Y., Shimozawa N., Orii T.;  
RT "Mucopolysaccharidosis IVA: identification of a common missense  
RT mutation 1113P in the N-Acetylgalactosamine-6-sulfate sulfatase  
RT gene.";  
RL Am. J. Hum. Genet. 57:556-563(1995).  
[5]  
RP VARIANTS MPS-IVA.  
RX MEDLINE=95315929; PubMed=7795586;  
RA Ogawa T., Tomatsu S., Fukuda S., Yamagishi A., Maruf Rezvi G.,  
RA Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.;  
RT "Mucopolysaccharidosis IVA: screening and identification of mutations  
RT of the N-acetylgalactosamine-6-sulfate sulfatase gene.";  
RL Hum. Mol. Genet. 4:341-349(1995).  
[6]  
RP VARIANTS MPS-IVA ARG-77; TRP-90; VAL-96; LEU-151; GLY-230 AND THR-291.  
RX MEDLINE=95359983; PubMed=7633425;  
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,  
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,  
RA Suzuki Y., Shimozawa N., Orii T.;  
RT "Mucopolysaccharidosis type IVA: identification of six novel  
RT mutations among non-Japanese patients.";  
RL Hum. Mol. Genet. 4:741-743(1995).  
[7]  
RP VARIANT MPS-IVA SER-487.  
RX MEDLINE=96047158; PubMed=7581409;  
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamada N., Isogai K.,  
RA Kato Z., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.;  
RT "Two new mutations, Q473X and N487S, in a Caucasian patient with  
RT mucopolysaccharidosis IVA (Morquio disease).";  
RL Hum. Mutat. 6:195-196(1995).  
[8]  
RP VARIANTS MPS-IVA ALA-138; SER-151 AND LEU-151.  
RX MEDLINE=96216146; PubMed=8651279;  
RA Tomatsu S., Fukuda S., Yamagishi A., Cooper A., Wraith J.E., Hori T.,  
RA Kato Z., Yamada N., Isogai K., Sukegawa K., Kondo N., Suzuki Y.,  
RA Shimozawa N., Orii T.;  
RT "Mucopolysaccharidosis IVA: four new exonic mutations in patients  
RT with N-acetylgalactosamine-6-sulfate sulfatase deficiency.";  
RL Am. J. Hum. Genet. 58:950-962(1996).  
[9]  
RP VARIANTS MPS-IVA CYS-94 AND VAL-97.  
RX MEDLINE=96423834; PubMed=8826435;  
RA Cole D.E.C., Fukuda S., Gordon B.A., Rip J.W., Lecouteur A.N.,  
RA Rupar C.A., Tomatsu S., Ogawa T., Sukegawa K., Orii T.;  
RT "Heteroallelic missense mutations of the galactosamine-6-sulfate  
RT sulfatase (GALNS) gene in a mild form of Morquio disease (MPS IVA).";

RL Am. J. Med. Genet. 63:558-565(1996).  
[10]  
RP VARIANTS MPS-IVA.  
RX MEDLINE=97442274; PubMed=9298823;  
RA Bunge S., Kleijer W.J., Tytki-Szymanska A., Steglich C., Beck M.,  
RA Tomatsu S., Fukuda S., Poorthuis B.J.H.M., Czartoryska B., Orii T.,  
RA Gal A.;  
RT "Identification of 31 novel mutations in the N-acetylgalactosamine-6-  
RT sulfatase gene reveals excessive allelic heterogeneity among patients  
RT with Morquio A syndrome.";  
RL Hum. Mutat. 10:223-232(1997).  
[11]  
RP VARIANTS MPS-IVA.  
RX MEDLINE=98041700; PubMed=9375852;  
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Ferreira P.,  
RA di Natale P., Tortora P., Fujimoto A., Kato Z., Yamada N., Isogai K.,  
RA Yamagishi A., Sukegawa K., Suzuki Y., Shimozawa N., Kondo N.,  
RA Sly W.S., Orii T.;  
RT "Fourteen novel mucopolysaccharidosis IVA producing mutations in  
RT GALNS gene.";  
RL Hum. Mutat. 10:368-375(1997).  
[12]  
RP VARIANTS MPS-IVA.  
RX MEDLINE=98180718; PubMed=9521421;  
RA Yamada N., Fukuda S., Tomatsu S., Muller V., Hopwood J.J., Nelson J.,  
RA Kato Z., Yamagishi A., Sukegawa K., Kondo N., Orii T.;  
RT "Molecular heterogeneity in mucopolysaccharidosis IVA in Australia  
RT and Northern Ireland: nine novel mutations including T312S, a common  
RT allele that confers a mild phenotype.";  
RL Hum. Mutat. 11:202-208(1998).  
[13]  
RP VARIANTS SER-393 AND MET-488.  
RX MEDLINE=98112415; PubMed=9452036;  
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamagishi A., Kato Z.,  
RA Yamada N., Isogai K., Sukegawa K., Suzuki Y., Shimozawa N., Kondo N.,  
RA Orii T.;  
RT "Fifteen polymorphisms in the N-acetylgalactosamine-6-sulfate  
RT sulfatase (GALNS) gene: diagnostic implications in Morquio disease.";  
RL Hum. Mutat. Suppl. 1:S42-S46(1998).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate groups of the N-  
CC acetyl-D-galactosamine 6-sulfate units of chondroitin sulfate and  
CC of the D-galactose 6-sulfate units of keratan sulfate.  
CC -1- SUBUNIT: OLIGOMER OF DISULFIDE LINKED 40- AND 15-KDA POLYPEPTIDES.  
CC -1- SUBCELLULAR LOCATION: Lysosomal.  
CC -1- DISEASE: Defects in GALNS are the cause of mucopolysaccharidosis  
CC type IVA (MPS-IVA) [MIM:233000]; also known as Morquio A syndrome.  
CC MPS-IVA is characterized by specific spondyloepiphyseal dysplasia,  
CC short trunk dwarfism, coxa valga, odontoid hypoplasia, corneal  
CC opacities, preservation of intelligence, and excessive urinary  
CC excretion of keratan sulfate and chondroitin-6-sulfate. Severely  
CC affected patients usually die of cardiopulmonary disturbance or  
CC cervical cord compression in the second or third decade of life.  
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; D17629; BAA04535.1; JOINED.  
CC EMBL; D17616; BAA04535.1; JOINED.  
CC EMBL; D17617; BAA04535.1; JOINED.  
CC EMBL; D17618; BAA04535.1; JOINED.  
CC EMBL; D17619; BAA04535.1; JOINED.  
CC EMBL; D17620; BAA04535.1; JOINED.  
CC EMBL; D17621; BAA04535.1; JOINED.  
CC EMBL; D17622; BAA04535.1; JOINED.  
CC EMBL; D17623; BAA04535.1; JOINED.  
CC EMBL; D17624; BAA04535.1; JOINED.  
CC EMBL; D17625; BAA04535.1; JOINED.



01-OCT-1996 (Rel. 34, Created)  
01-OCT-1996 (Rel. 34, Last sequence update)  
15-SEP-2003 (Rel. 42, Last annotation update)  
Arylsulfatase A precursor (EC 3.1.6.8) (ASA) (Cerebroside-sulfatase).  
ARSA OR AS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
RP STRAIN=129/SV, and C57BL/6J;  
RX MEDLINE=94245194; PubMed=7910580;  
RA Kreyling J., Polten A., Hess B., von Figura K., Menz K., Steiner F.,  
RA Gieselmann V.,  
RA "Structure of the mouse arylsulfatase A gene and cDNA.";  
RL Genomics 19:249-256(1994).  
[2]  
SEQUENCE OF 32-66 FROM N.A.  
RP MEDLINE=92241876; PubMed=1572648;  
RA Grompe M., Pieretti M., Caskey C.T., Ballabio A.;  
RX "The sulfatase gene family: cross-species PCR cloning using the MOPAC  
RT technique.";  
RL Genomics 12:755-760(1992).  
CC -1- FUNCTION: Hydrolyzes cerebroside sulfate.  
CC -1- CATALYTIC ACTIVITY: A cerebroside 3-sulfate + H(2)O = a  
CC cerebroside + sulfate.  
CC -1- SUBUNIT: EXISTS BOTH AS A SINGLE CHAIN OF 58 kDa (COMPONENT A)  
CC OR AS A CHAIN OF 50 kDa (COMPONENT B) LINKED BY DISULFIDE BOND(S)  
CC TO A 7 kDa CHAIN (COMPONENT C).  
CC -1- SUBCELLULAR LOCATION: Lysosomal.  
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
CC EMBL: X73231; CAA51703.1; -  
CC EMBL: X73230; CAA51702.1; -  
CC EMBL: M82876; AAA37260.1; -  
CC PIR: A54190; A54190.  
CC HSSP: P15289; LAUK.  
CC MGD: MGI:88077; Arsa.  
CC GO: GO:0000299; C: integral membrane protein of membrane fraction; IDA.  
CC GO: GO:0005886; C: plasma membrane; IDA.  
CC GO: GO:0007339; P: binding of sperm to zona pellucida; IMP.  
CC InterPro: IPR000917; Sulfatase.  
CC Pfam: PF00884; Sulfatase; 1.  
CC PROSITE: PS00523; SULFATASE\_1; 1.  
CC PROSITE: PS00149; SULFATASE\_2; 1.  
CC Hydrolase; Signal; Glycoprotein; Lysosome.  
KW SIGNAL  
FT 1 17  
FT CHAIN 18 506  
FT MOD\_RES 68 68  
FT 2-AMINO-3-OXOPROPIONIC ACID (BY  
SIMILARITY).  
FT BY SIMILARITY.  
FT BY SIMILARITY.  
FT BY SIMILARITY.  
FT BY SIMILARITY.  
FT BY SIMILARITY.  
FT BY SIMILARITY.  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 349 349  
FT CARBOHYD 506 AA; 53776 MW; 74BSL6401CF18DF6 CRC64;  
SEQUENCE  
Query Match 4.4%; Score 210.5; DB 1; Length 506;  
Best Local Similarity 25.2%; Pred. No. 1.1e-07;  
Matches 94; Conservative 64; Mismatches 158; Indels 57; Gaps 18;



-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; J052225; AAA51784.1; -  
EMBL; M32373; AAA51779.1; -  
EMBL; X72735; CAAS1272.1; -  
EMBL; X72736; CAAS1272.1; JOINED.  
EMBL; X72737; CAAS1272.1; JOINED.  
EMBL; X72738; CAAS1272.1; JOINED.  
EMBL; X72739; CAAS1272.1; JOINED.  
EMBL; X72740; CAAS1272.1; JOINED.  
EMBL; X72741; CAAS1272.1; JOINED.  
EMBL; X72742; CAAS1272.1; JOINED.  
EMBL; X72743; CAAS1272.1; JOINED.  
EMBL; S57777; AAB19988.1; -  
PIR; S35990; KJHUAB.  
PDB; 1FSU; 04-FEB-98.  
Genew; HGNC:714; ARSB.  
MIM; 253200; -  
MIM; 272200; -  
GO; GO:0005764; C-lysosome; TAS.  
GO; GO:0004065; Fatty acid metabolism; TAS.  
GO; GO:0007041; P-lysosomal transport; TAS.  
InterPro; IPR000917; Sulfatase.  
Pfam; PF00884; Sulfatase; 1.  
PROSITE; PS00523; SULFATASE 1; 1.  
PROSITE; PS00149; SULFATASE 2; 1.  
Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis;  
Disease mutation; Polymorphism; 3D-structure.  
SIGNAL 1 36  
CHAIN 37 533  
ACT SITE 147 147  
MOD RES 91 91  
DISULFID 117 521  
DISULFID 121 155  
DISULFID 181 192  
DISULFID 405 447  
CARBOHYD 188 188  
CARBOHYD 279 279  
CARBOHYD 291 291  
CARBOHYD 366 366  
CARBOHYD 426 426  
CARBOHYD 458 458  
VARIANT 92 92  
VARIANT 95 95  
VARIANT 117 117  
VARIANT 137 137  
VARIANT 152 152  
VARIANT 160 160  
VARIANT 210 210  
VARIANT 236 236  
VARIANT 302 302  
VARIANT 376 376  
VARIANT 393 393  
VARIANT 405 405  
VARIANT 498 498  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .).  
N-LINKED (GLCNAC. . .) (PROBABLE).  
N-LINKED (GLCNAC. . .).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
T -> M (IN MPS-VI; MILD FORM).  
R -> Q (IN MPS-VI; MILD/SEVERE FORM).  
C -> R (IN MPS-VI; SEVERE FORM).  
G -> V (IN MPS-VI; INTERMEDIATE FORM).  
R -> W (IN MPS-VI; INTERMEDIATE FORM).  
R -> Q (IN MPS-VI; INTERMEDIATE FORM).  
Y -> C (IN MPS-VI; MILD/INTERMEDIATE).  
L -> P (IN MPS-VI; MILD FORM).  
G -> R (IN MPS-VI; SEVERE FORM).  
V -> M.  
H -> P (IN MPS-VI; MILD/SEVERE FORM).  
C -> Y (IN MPS-VI; MILD FORM).  
L -> P (IN MPS-VI; MILD/SEVERE FORM).

FT CONFLICT 358 358 /FTId=VAR 007306.  
FT STRAND 46 52 V -> M (IN REF. 3).  
FT TURN 57 58  
FT HELIX 61 63  
FT TURN 64 64  
FT TURN 70 77  
FT TURN 78 78  
FT STRAND 80 82  
FT TURN 83 84  
FT STRAND 85 86  
FT HELIX 93 100  
FT HELIX 104 107  
  
Query Match 4.4%; Score 210.5; DB 1; Length 533;  
Best Local Similarity 20.8%; Pred. No. 1.2e-07;  
Matches 102; Conservative 82; Mismatches 173; Indels 133; Gaps 23;  
  
Qy 2 GPESLVLCLSATVPSILL--GGSSAFLSHRLKGRFQDRNRNIRENIIILVLTDD--QD 55  
Db 14 GPRRLLPVPLPILLLLAPPQSGAGSR-----PHELVLADLGLWND 59  
Qy 56 VELGSMQVWVKTRIMEQGAHFINAFVTPMCPSPRSSILTQY-----VHNHNTYNN 110  
Db 60 VGFHSGKIRTFHDLAAGGV-LLDNYYTOPLCTPSRSOLLTGRYQIRGTGLQHIIWPCQ 118  
Qy 111 ENCSSPSWQACHESRTFAYVILNSTGYRTAFPGK-YLNEYNGSVVP--GWKEWVGLKNS 167  
Db 119 PSC-----VPLDEKLLPOLKEAGYTTMVGKHLGMYRKCLPFRGFDYTFVGLLGS 172  
Qy 168 RFY-----NVTLC-----RNGVKEKHSYSDYLDLTIDNDSVSFFRTSKMY 211  
Db 173 EDYSHERCCTLIDALNVTRCALDFRDG--EEVATGYKNWYSNIFTKRALITNHP-- 227  
Qy 212 PRRPVLWVSHAAHPGPDSPQYSLFPNASCHITPSPYAPNPDKHMYGTGPMKEI 271  
Db 228 PERFLFLYLALQSVHEPLQVPEYLK-----PYDFIQDKNRH---HYAG----- 268  
Qy 272 HMEFTNMQRLQTLMSVDDSMETIYNMLVETGELDNTIYVYVADHGYHIGQFGLVKGK 331  
Db 269 -----MVSLM--DEAVGNTAALKSGSLWNTVFIFSTDNG--GQ-TLAGGN 310  
Qy 332 SMP-----YEFDIRVPFYVGRPNV-EAGCLNPHIVLNIDLAFTILDIA--GLDIPAD 380  
Db 311 NWPLRGRKWSLWEGGVGVGFVASFLLKQKQVKNRELIHISDWLPTLVKLARGHTNGTKE 370  
Qy 381 MDGKSLKLLDTERPVRNRFHLKCKKRVWEDSFIVERGKLLHKEDNDKVD----- 429  
Db 371 LDGFDVWKTISEGSPSPRI-----ELRNIDPNFVDSPPCRNSMAP 412  
Qy 430 AQEENFLPKY 439  
Db 413 AKDDSSILPEY 422

Search completed: February 14, 2004, 22:54:56  
Job time : 44 secs